PLOS ONE

RESEARCH ARTICLE

Crystal structure of the kinase domain of a receptor tyrosine kinase from a choanoflagellate, *Monosiga brevicollis*

Teena Bajaj[1], John Kuriyan[1,2,3,4,5¤], Christine L. Gee[2,4,5]*

1 Graduate Program in Comparative Biochemistry, University of California, Berkeley, Berkeley, California, United States of America, **2** Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California, United States of America, **3** Department of Chemistry, University of California, Berkeley, Berkeley, California, United States of America, **4** Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, California, United States of America, **5** California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, California, United States of America

¤ Current address: Departments of Biochemistry and Chemistry, Vanderbilt University, Nashville, Tennessee, United States of America
* christinelgee@berkeley.edu

Citation: Bajaj T, Kuriyan J, Gee CL (2023) Crystal structure of the kinase domain of a receptor tyrosine kinase from a choanoflagellate, *Monosiga brevicollis*. PLoS ONE 18(6): e0276413. https://doi.org/10.1371/journal.pone.0276413

Editor: Matteo De March, University of Nova Gorica, SLOVENIA

Received: October 5, 2022

Accepted: April 28, 2023

Published: June 13, 2023

Data Availability Statement: The PDB file and structure factors for the structure reported in this paper are available from the RCSB database (PDB ID 8E4T) doi 10.2210/pdb8e4t/pdb

Funding: This work was supported by the Howard Hughes Medical Institute (https://www.hhmi.org/) JK is an HHMI Investigator. Beamline 8.2.1, run by the Berkeley Center for Structural Biology at the Advanced Light Source is a is supported by the Howard Hughes Medical Institute and by the Department of Energy Office of Science User Facility under Contract No. DE-AC02-05CH11231

Abstract

Genomic analysis of the unicellular choanoflagellate, *Monosiga brevicollis* (MB), revealed the remarkable presence of cell signaling and adhesion protein domains that are characteristically associated with metazoans. Strikingly, receptor tyrosine kinases, one of the most critical elements of signal transduction and communication in metazoans, are present in choanoflagellates. We determined the crystal structure at 1.95 Å resolution of the kinase domain of the *M. brevicollis* receptor tyrosine kinase C8 (RTKC8, a member of the choanoflagellate receptor tyrosine kinase C family) bound to the kinase inhibitor staurospaurine. The chonanoflagellate kinase domain is closely related in sequence to mammalian tyrosine kinases (~ 40% sequence identity to the human Ephrin kinase domain EphA3) and, as expected, has the canonical protein kinase fold. The kinase is structurally most similar to human Ephrin (EphA5), even though the extracellular sensor domain is completely different from that of Ephrin. The RTKC8 kinase domain is in an active conformation, with two staurosporine molecules bound to the kinase, one at the active site and another at the peptide-substrate binding site. To our knowledge this is the first example of staurospaurine binding in the Aurora A activation segment (AAS). We also show that the RTKC8 kinase domain can phosphorylate tyrosine residues in peptides from its C-terminal tail segment which is presumably the mechanism by which it transmits the extracellular stimuli to alter cellular function.

Introduction

Receptor tyrosine kinases are crucial signaling molecules in animal cells, regulating inter-cellular responses to extracellar signals via numerous different pathways dependent on the

The funders has no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared no competing interests exist.

organism involved. These receptors consist of an extracellular ligand-binding domain, a single transmembrane helix, and a cytoplasmic tyrosine kinase domain [1, 2]. Binding of ligands to the extracellular domain of the receptor leads to dimerization and activation of the kinase domain. The activated kinase domain then autophosphorylates tyrosine residues present in the activation loop in the kinase domain, thus sustaining or amplifying the activation. Other sites of phosphorylation in the C-terminal tail of the receptor usually are involved in recruitment of further components of the signaling pathway. In some instances, loops in the kinase domain, or in the juxtamembrane segment bridging the transmembrane helix and the kinase domain are phosphorylated to alter the kinase activity. Thus, in summary, phosphorylated tyrosine residues switch on kinase activity, or act as docking sites for proteins that transmit the signal further downstream [3–5].

The discovery of numerous genes encoding receptor tyrosine kinases in choanoflagellates was remarkable [6–10]. Phylogenetic studies suggest that choanoflagellates are the closest living relatives to metazoans, indicating that receptor tyrosine kinases were present in the common ancestor of choanoflagellates and metazoans. The choanoflagellate *M. brevicollis* contains 128 protein kinases, including 88 receptor tyrosine kinases [11]. These choanoflagellate receptor tyrosine kinases are classified into 15 families (denoted A through M, FGTK and LRTK) [11]. They are not direct orthologs of metazoan receptor tyrosine kinases as they have extracellular domains that appear unrelated to those of metazoan receptor tyrosine kinases [11]. The structural and functional aspects of choanoflagellate receptor tyrosine kinases are still relatively unexplored.

The kinase described in this study is from the *M. brevicollis* receptor tyrosine kinase C (RTKC) family. This family has 10 members (denoted RTKC1 to RTKC10). The RTKC family is a distinct clade of kinases and shares a common ancestor with the Ephrin kinases but does not share all of the Ephrin-specific conserved motifs [12]. Most of the members of this family contain Cys-rich and Hyalin-rich (HYR) extracellular domains, a transmembrane domain, a kinase domain and a C-terminal tail with a CAP-Gly domain. The HYR domain is predicted to be related to immunoglobulin (Ig) and fibronectin type3 (FN3) domains in metazoans, and might be involved in cell adhesion [11, 13]. The metazoan CAP-Gly domains bind to microtubules [14]. The presence of a CAP-Gly domain in *M. brevicollis* C family receptor tyrosine kinases therefore suggests that these kinases may interact with the cytoskeleton [11]. The *M. brevicollis* RTKC8 gene codes for a 2031 amino acid protein, It consists of an extracellular domain of ~1487 residues containing two Cys-rich domains and four HYR domains, a transmembrane domain (TM) domain of ~19 residues and an intracellular domain consisting of a juxtamembrane domain (JM, ~37 residues) a kinase domain (~254 residues) and a C-terminal tail containing a 40 residue CAP-Gly domain and 10 tyrosine residues (Fig 1).

We determined the crystal structure of the RTKC8 kinase domain to 1.95 Å resolution. The RTKC8 kinase domain is co-crystallized with the kinase inhibitor staurospaurine in two positions, one the canonical active site position and one in the peptide binding site or Aurora A activation segment (AAS). We also demonstrated the kinase is active on tyrosine residue containing peptides derived from its C-terminal tail segment.

Materials and methods

Cloning of the RTKC8 kinase domain

The sequence of the RTKC8 kinase matches that reported in the Kinase.com database (http://kinase.com/kinbase) [11]. The DNA (NCBI accession id: XM_001750163.1), encoding for the RTKC8 kinase domain (1543–1796) flanked by a 13 residues of the preceeding sequence and 15 residues of the following sequence, residues 1530–1811, was amplified using standard PCR

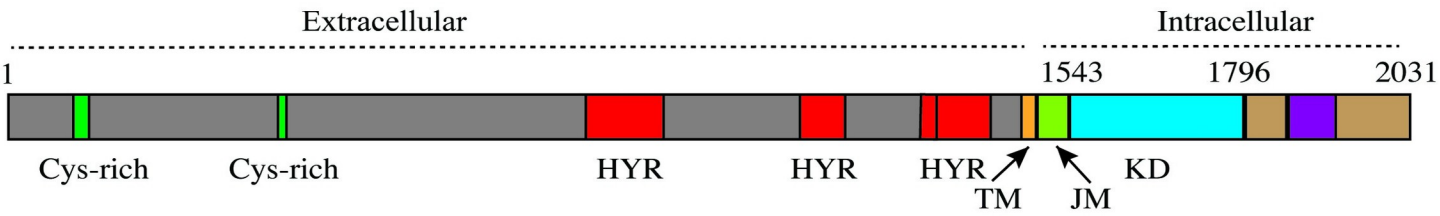

Fig 1. Architecture of the RTKC8 receptor tyrosine kinase and purification of its kinase domain. Domain organization of full length RTKC8, consisting of Cys-rich and hyalin rich (HYR) domains in the extracellular domain, followed by a transmembrane domain (TM) and an intracellular domain. The intracellular domain contains a juxtamembrane (JM) domain, a kinase domain, and a C-terminal tail with tyrosine residues (red circles) and a CAP-Gly domain.

https://doi.org/10.1371/journal.pone.0276413.g001

protocols from a *Monosiga* cDNA library using the forward primer 5′AGGGCCATATGCA GCTTTCCAAGGAGCCGCGCG3′ and the reverse primer 5′CCTTTGAATTCTCAGCTTT GGAAGAGGCTGGACATGTCCGAATCAGT3′. The PCR product size was confirmed using agarose gel electrophoresis. The amplified DNA was ligated into the pFastbac1 insect-cell expression vector using EcoRI (5') and NdeI (3') restriction enzyme sites. The ligated product was transformed into *E. coli* Top10 chemically competent cells and was selected with Ampicillin on LB Agar. Next day, overnight cultures were set up from the Top10 bacterial colonies to isolate the plasmid and the plasmid was sequenced to confirm the cloned construct. The expressed construct consisted of an N-terminal hexahistidine tag and a TEV protease site followed by the *M. brevicollis* TRKC8 core kinase domain with short flanking sections from the RTKC8 protein sequence either side (S1A Fig).

Expression of the RTKC8 kinase domain

The pFastbac1-RTKC8-kinase vector carrying the ampicillin resistance gene was transformed into DH10bac *E. coli* cells for transposition into the bacmid. The outgrowth (diluted 1:100) was plated over Luria-Bertani (LB) agar plates containing kanamycin (50 μg/mL), gentamycin (7 μg/mL), tetracycline (10 μg/mL), isopropyl β-D-1-thiogalactopyranoside (IPTG) (40 μg/mL), bluo-gal (100–300 μg/mL) and incubated at 37˚C. After 24–48 hours, the white colonies (containing the cloned construct) were cultured overnight to isolate the bacmid. The bacmid was transfected into *Spodoptera frugiperda* (Sf9) cells to prepare baculovirus stock, which was further amplified via three cycles of infection. After 3 cycles of virus amplification, the supernatant of this third passage (P3 virus) was used to infect another 2 liter batch of Sf9 cells grown to 2.5 x 106 cells/ mL in the presence of 10% antibiotics (penicillin and streptomycin). Sf9 cells were harvested after 48 hours. Each liter of culture was resuspended in 20 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 20% (v/v) glycerol, 10 mM $CaCl_2$, 10 mM $MgCl_2$, 50 mM imidazole, 0.5 mM TCEP) containing 10 μg/mL DNaseI and a protease inhibitor cocktail of 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) (0.2 mM), benzamidine (0.5 mM) and leupeptin (0.005 mM). The resuspended lysate was stored at -80˚C until thawed for purification.

Purification of the RTKC8 kinase domain

To purify the protein, sf9 cells were thawed and lysed by passing three times through an Avestin EmulsiFlex C50 cell homogenizer with 1500 psi pressure at 4˚C. The suspension was clarified by ultracentrifugation at 40,000 rpm for 1 hour at 4˚C. The supernatant was filtered using a 0.45 μM syringe filter, loaded onto a Ni-NTA affinity column (GE Healthcare Life Sciences) pre-equilibrated with Buffer A (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 25 mM imidazole, 0.5 mM TCEP) and the flow-through was collected. The resin was washed with

20 column volumes of Buffer A to remove the non-specific proteins. Then, the RTKC8 protein was eluted in Buffer B (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, 250 mM imidazole, 0.5 mM TCEP). To cleave the His-tag, TEV protease was added to the pooled fractions containing the protein and dialyzed into Buffer C (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, 0.5 mM TCEP) overnight at 4˚C. A subtractive Ni-NTA purification was run to remove TEV protease and other impurities. The flow through was collected, concentrated and purified on a Superdex 200 16/600 column (GE Healthcare Life Sciences) in Buffer C. Fractions containing RTKC8 kinase were either pooled and concentrated to 105 μM for biochemical experiments and stored at -80˚C or concentrated to 200 μM and used immediately for crystallization experiments. The purification yield was ~1 mg of purified RTKC8 kinase from 1 liter of Sf9 insect cells. The elution volume from size exclusion was consistent with the molecular weight of a monomeric protein (28.5 KDa) (S1B Fig) and ran as a single chromatography band on SDS PAGE (S1C Fig).

Phosphorylation and dephosphorylation analysis of RTKC8 kinase domain

To determine the phosphorylation state of the purified RTKC8 kinase domain, the protein was probed using an anti-phosphotyrosine antibody (4G10) by western blot. The protein was run on 12% (w/v) acrylamide SDS-PAGE gel at a final concentration of 1 μM and transferred to polyvinylidene difluoride (PVDF) membrane (Millipore) using a semi-dry transfer apparatus. Protein samples were transferred using Towbin transfer buffer (25 mM tris, 192 mM glycine, 20% (v/v) methanol). The membrane was blocked using blocking solution (5% (w/v) BSA dissolved in Tris-buffered saline (20 mM Tris pH 7.5, 150 mM NaCl), 0.1% (v/v) Triton X-100). The membrane was probed using anti-phosphotyrosine 4G10 primary antibody (Millipore, 05–321) (1:3000), diluted in blocking solution overnight at 4˚C. Next day, the blot was washed with 1X Tris-buffered saline containing 0.1% (v/v) Triton X-100 four times for 5 minutes each with gentle shaking at room temperature. The blot was probed using anti-Mouse IgG secondary antibody linked with horse radish peroxidase (Cell Signaling Technology, #7076) (1:3000) for 1 hour at room temperature on shaker. The blot was washed again with 1X Tris-buffered saline solution containing 0.1% (v/v) Triton X-100 four times for 5 minutes each with gentle shaking at room temperature. The blot was developed using the Western Bright kit (Advansta) and imaged using a Bio-Rad Gel Doc.

For dephosphorylation assays, the RTKC8 kinase domain and YopH phosphatase, each at a final concentration of 1 μM, were incubated together at room temperature and the dephosphorylation reaction was stopped at different time intervals by adding SDS-buffer. Phosphotyrosine levels were detected by western blot as described above.

In vitro kinase activity assays

Kinase activity measurements were carried out using a coupled kinase assay (Pyruvate Kinase/Lactate Dehydrogenase (PK/LDH) ATPase assay) in presence of different substrate peptides [15, 16]. This assay measures ADP production, which is coupled to NADH oxidation. In this experiment, the reaction solution contained 50 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM MgCl2, 1 mM phosphoenolpyruvate (PEP), 300 mg/mL NADH, 2 mM sodium orthovanadate (Na_3VO_4), 100 μM ATP, and an excess of pyruvate kinase and lactate dehydrogenase (approximately 120 units/mL and 80 units/mL, respectively, added from a commercially available mixture of the two enzymes). Human c-Src kinase domain phosphorylation of a preferred peptide substrate, PKCδ-Y313 (SSEPVGI**Y**QGFEKKT) and a non-preferred peptide, PDGFRα-Y762 (SDIQRSL**Y**DRPASAK), in which the Tyr 768 in the wildtype PDGFRα sequence was mutated to alanine in the peptide so as to present a single substrate tyrosine residue, were used as

controls in these kinase assays [17]. The tyrosine phosphorylation kinetics for RTKC8 with these two previously mentioned peptides and three peptides from the RTKC8 C-terminal tail (purchased from Elim Biopharmaceuticals, Inc.), corresponding to the residues surrounding Tyr 1828 (EPETDEV**Y**GNEGNVS), Tyr 1897 (GTVGEHE**Y**FDCADQH), and Tyr 1990 (ASDNELL**Y**DMGRAEA) were measured. The reactions were initiated by the addition of either c-Src at 1 μM or RTKC8 kinase domain at 5 μM, to 250 mM peptide and were conducted in duplicate. The reaction progress was monitored by measuring absorbance at 340 nm every 30 s at 25˚C on a SpectraMax plate reader.

Crystallization and data collection

The freshly purified RTKC8 kinase domain (10–15 mg/mL, in Buffer C) was mixed with a final concentration of 1 mM staurosporine (10 mM stock, dissolved in 100% (v/v) DMSO) and incubated on rotator for 2 hours at room temperature. The crystallization trays (96 well) were set up using sitting-drop vapor diffusion with a Mosquito crystallization robot (TTP Labtech) using commercial sparse matrix screens. 0.1 μL of the protein was mixed with 0.1 μL of the reservoir solution. The drops were equilibrated against 50 μL of reservoir volume and the trays were incubated at 20˚C. A condition from the PACT screen (Qiagen) yielded crystals suitable for diffraction (0.1 M succinic acid, sodium dihydrogen phosphate and glycine (SPG) buffer pH 6.0, 25% (w/v) polyethylene glycol (PEG) 1500 [18]. The crystals directly from the 96 well screen were cryoprotected with a solution of the well condition with 20% (v/v) glycerol and X-ray diffraction data were collected at Lawrence Berkeley National Laboratory Advanced Light Source, Beamline 8.2.1 with 0.9998 Å wavelength X-rays.

Structure determination

The reflections were indexed and integrated by XDS [19, 20]. The data were scaled and merged with Pointless [21] and Aimless [22], in the CCP4 suite [23]. The structure was solved by molecular replacement with Phaser-MR [24] using the structure of tyrosine kinase AS—a hypothetical protein with a sequence corresponding to a common ancestor of Src and Abl with ~ 43% identity to the RTKC8 kinase domain (PDB 4UEU [25]) as a search model. The structure was rebuilt using Phenix-AutoBuild and refined with multiple cycles of refinement with phenix.refine [26] and model building with Coot [27]. We used Molprobity [28] statistics provided in the Phenix suite to validate the structure. The calculated difference Fourier map (Fo-Fc) revealed clear electron density for two molecules of bound staurosporine per kinase domain and one phosphate, resumed to be from the crystallization buffer. The phosphate (PO4) and staurospaurine (STU) ligands were refined with full occupancy using the standard cif dictionaries included in the phenix suite. The B-factors for the ligands were comparable to the B-factors of the protein amino acid residues co-ordinated with them. The coordinates and structure factors of the final structure have been deposited in the Protein Data Bank (PDB) with the PDB ID 8E4T. Software used in this project was curated by SBGrid [29]. Figures depicting the structure were generated with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

Results

RTKC8 kinase domain is phosphorylated

Western blot analysis was performed to check the phosphorylation status of the purified RTKC8 kinase. The immunoblot was probed with an anti-phosphotyrosine antibody and showed that the purified RTKC8 kinase had at least one phosphorylated tyrosine residue

which could be dephosphorylated by the tyrosine-protein phosphatase YopH (Fig 2A). The band intensity was decreased to ~30% of the initial level after five minutes and to less than ~5% after an hour.

Substrate specificity of the RTKC8 kinase

The phosphorylation rates of the peptides human c-Src preferred peptide (PKCδ-Y313) and non-preferred peptide (PDGFRα-Y762) [17] by c-Src (1 μM) and RTKC8 kinase domain (5 μM) were assayed (Fig 2B and 2C). RTKC8 kinase appears to have a different substrate specificity than c-Src, showing more rapid phosphorylation of PDGFRα-Y762 than PKCδ-Y313. Kinase assays were also performed with three RTKC8 tail peptides comprised of the sequences surrounding Tyr 1828, Tyr Y1897, or Tyr Y1990. These experiments showed that the different tail peptides were phosphorylated at different rates. The RTKC8 peptide containing Tyr 1990 had the highest rate of phosphorylation by RTKC8 kinase (Fig 2B and 2D). Both the c-Src non-preferred peptide (PDGFRα-Y762) and the tail peptide (RTKC8-Y1990) have Leu before the Tyr and Asp after the phosphosite tyrosine residue. This result suggests that the preferred substrates of the RTKC8 kinase contain a hydrophobic residue at the position immediately before the phosphosite and an acidic residue immediately after it.

Structure determination

Crystallization trials of the RTKC8 kinase domain with staurospaurine yielded triangular shaped crystals, 70–100 μm in diameter. The crystals were of the orthorhombic space group $P2_12_12$ with cell dimensions a = 81.1, b = 55.5, c = 60.6 (Å) with a solvent content of 46.91% and a Matthews coefficient of 2.32 and one molecule in the asymmetric unit. The crystal structure of the kinase domain was determined to 1.95 Å resolution and the model was built with a final Rwork/Rfree (%) of 20.4/25.5 (Table 1). The crystallographic model includes residues 1532 to 1804, two staurosporine molecules, 100 water molecules and one phosphate group (Fig 3A). The first three residues (1530–1531), the last seven residues (1805–1811) and a region of the activation loop (1694–1698) in the expressed protein construct were not modeled as there was no interpretable electron density for these residues. The tyrosine residues which would be expected to be phosphorylated in the activation loop (Tyr 1697 and Tyr 1698) were in the disordered activation loop region. The staurosporine inhibitor bound to the kinase domain at two sites, one in the ATP binding site (Fig 4A) and other on the C-lobe in the peptide-substrate binding site (Figs 3C and 4B).

The choanoflagellate RTKC8 kinase domain crystallized in an active conformation

The choanoflagellate kinase domain exhibits the conserved structural features of a typical kinase domain [30, 31] with two lobes, the N-lobe and C-lobe. The activation loop is a stretch of 29 amino acids in the C-lobe, beginning with the conserved Asp-Phe-Gly (DFG) motif and ending at Ala-Pro-Glu (APE). The activation loop of the RTKC8 kinase domain contains two consecutive tyrosine residues for which there is no observable electron density.

The RTKC8 kinase crystallized in an active conformation with a RMSD of 1.0 Å over 201 Cα atoms to the structure of Lck in an active conformation [32]. The conserved Glu 1592 of the αC-helix is rotated towards the active site to form a salt bridge with Lys 1575 of the β3-strand, oriented as "αC-helix-in". The distance between αC-helix-Glu-Cβ and β3-Lys-Cβ atoms is 8.1 Å (C-helix-in and C-helix-out structures are defined with the distance of $\leq$10 Å or >10 Å, respectively) [33]. The kinase is crystallized in "DFGin" conformation where the DFG-Asp is pointed towards the active site and the DFG-Phe interacts with the αC-helix. The

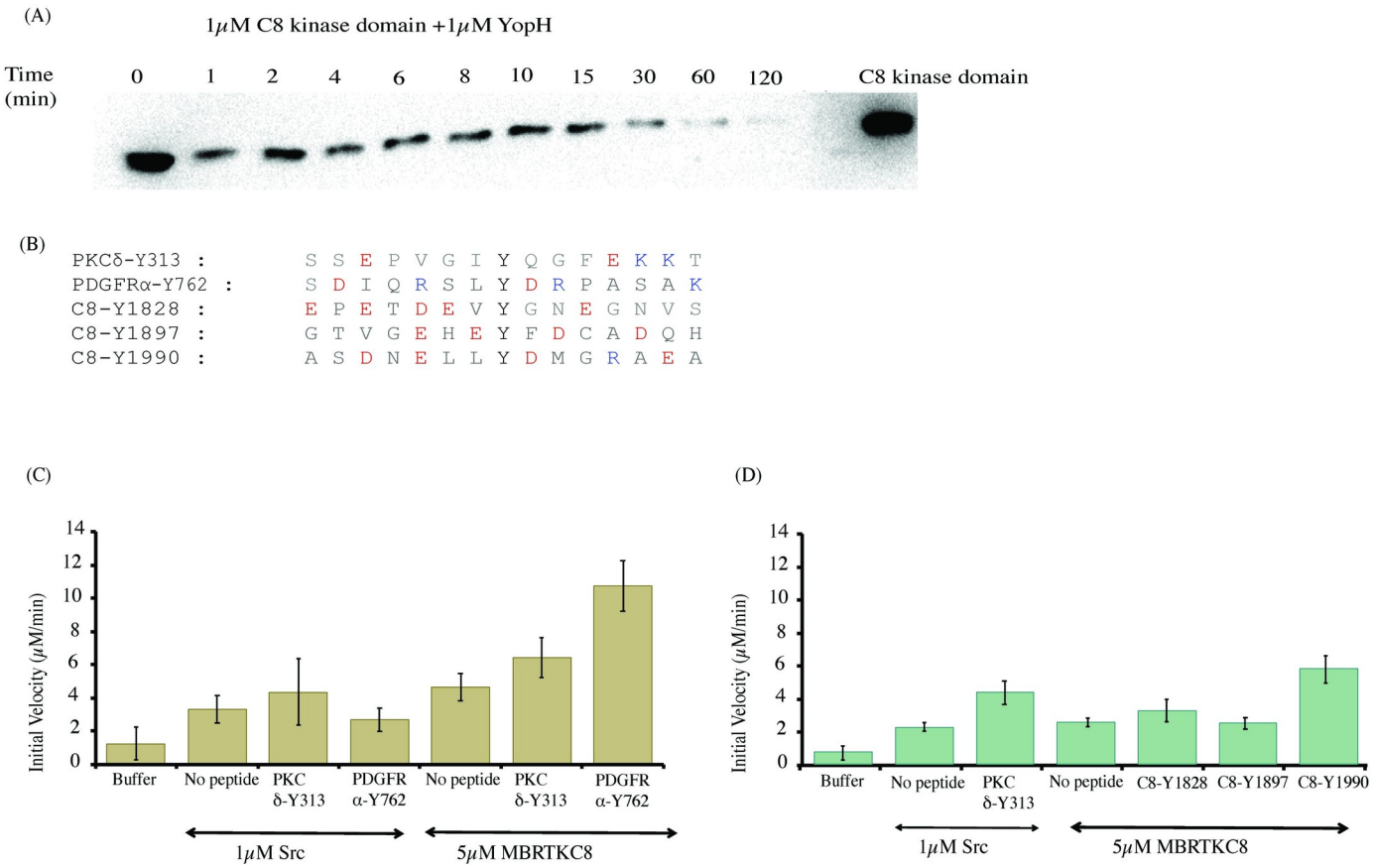

Fig 2. Phosphorylated RTKC8 kinase domain is active and phosphorylates tyrosine residues in tail segment peptides. (A) The RTKC8 kinase domain was dephosphorylated by YopH phosphatase and samples were taken at different time intervals and probed with an anti-phosphotyrosine antibody. The control showed that the band intensity on the western blot for the kinase domain without YopH is unchanged over 2 hours at room temperature while the band intensity decreased rapidly upon the addition of the tyrosine phosphatase YopH. (B) The sequence of tyrosine phosphosite peptides (PKCδ-Y313, PDGFRα-Y762, RTKC8-Y1828, RTKC8-Y1897, RTKC8-Y1990) used as substrates for the kinase (C) Phosphorylation activity of RTKC8 kinase domain on the tyrosine containing peptides PKCδ-Y313 and PDGFRα-Y762 (D) Phosphorylation activity of RTKC8 kinase on tyrosine residue containing peptides from the tail segment (RTKC8-Y1828, RTKC8-Y1897, RTKC8-Y1990). Kinase assays were performed in duplicate.

https://doi.org/10.1371/journal.pone.0276413.g002

Asp backbone and its sidechain is oriented in such a manner that the carbonyl group of the Ala prior to the DFG motif forms a hydrogen bond with the nitrogen of the histidine sidechain in HRD motif [33]. The distance between the sidechain nitrogen of His 1664 and the carbonyl group of the Ala 1683 is 3.1 Å, consistent with formation of a hydrogen bond between them. The activation loop is extended away from the C-lobe of kinase domain. In this extended conformation, the cleft is accessible to ATP and peptide substrates [34, 35]. The orientation of the activation loop leads to the formation of a hydrogen bond between the backbone N atom of the sixth residue in the loop (DFGxx**X**) and the backbone carbonyl of the residue prior to the HRD motif in the catalytic loop. The amide nitrogen of Arg 1689 (DFGxx**R**) and the carbonyl of Ile 1663 form a hydrogen bond of 3 Å. Another major structural feature of this conformation is the presence of type I β-turn at the Phe in the DFG motif, which forms a hydrogen bond between the carboxyl group of the Phe and the NH of the Ala two residues after the DFG motif [33]. The hydrogen bond length between Phe 1685 and Ala 1688 is 3.2 Å. This interaction plays an important role in orientating the activation loop to an extended conformation. All these interactions and structural features present the signature of the active conformation of the kinase domain (Fig 3C).

Table 1. Structure determination and refinement of RTKC8 kinase domain bound to staurosporine.

Data Collection	
Wavelength (Å)	0.9998
Space group	$P2_12_12$
Cell dimensions (Å)	a = 81.1, b = 55.5, c = 60.6
Resolution (Å)	48.55–1.95 (2.0–1.95)
Rmerge	0.23 (3.65)
Rmeas	0.23 (3.76)
Rpim	0.05 (0.88)
Total Observations	539294 (23983)
Unique Observations	20593 (1414)
CC1/2	0.99 (0.45)
Completeness (%)	99.9 (99.3)
Multiciplicity	26.2 (17.0)
I/σ (I)	12.4 (0.9)
Refinement Statistics	
No. reflections in Rfree test set	1039 (142)
Resolution	48.55–1.95 (2.05–1.95)
R-work (%)	20.4 (29.0)
Rfree (%)	25.5 (32.0)
No. of Atoms	
Protein	2111
Ligand	75
STU	70
PO_4	5
Water	100
Average B-factor ($Å^2$)	
Proteins	35.07
Ligand	38.54
Solvent	36.37
Ramachandran Statistics	
Favoured (%)	97.35
Disallowed (%)	0.00
Root mean square deviation from ideality	
Bonds (Å)	0.007
Angle (˚)	0.847
Molprobity score	1.23
Molprobity clash score	2.08
Rotamer outliers (%)	1.76
RSCC ligands	
STU1 (active site)	0.95
STU2 (AAS site)	0.81
PO4	0.97

https://doi.org/10.1371/journal.pone.0276413.t001

Staurosporine interactions with the RTKC8 kinase domain

Two molecules of staurosporine are bound to the RTKC8 kinase. One is in the ATP-binding site as seen in several other structures of kinase:staurospaurine structures (eg. CDK2 (1AQ1 [36]) and ZAP-70 (1U59 [37]) (Fig 3A), and the other is bound in the peptide substrate binding site in the C-lobe (Fig 3B and 3C). The active site inhibitor is positioned in the large groove

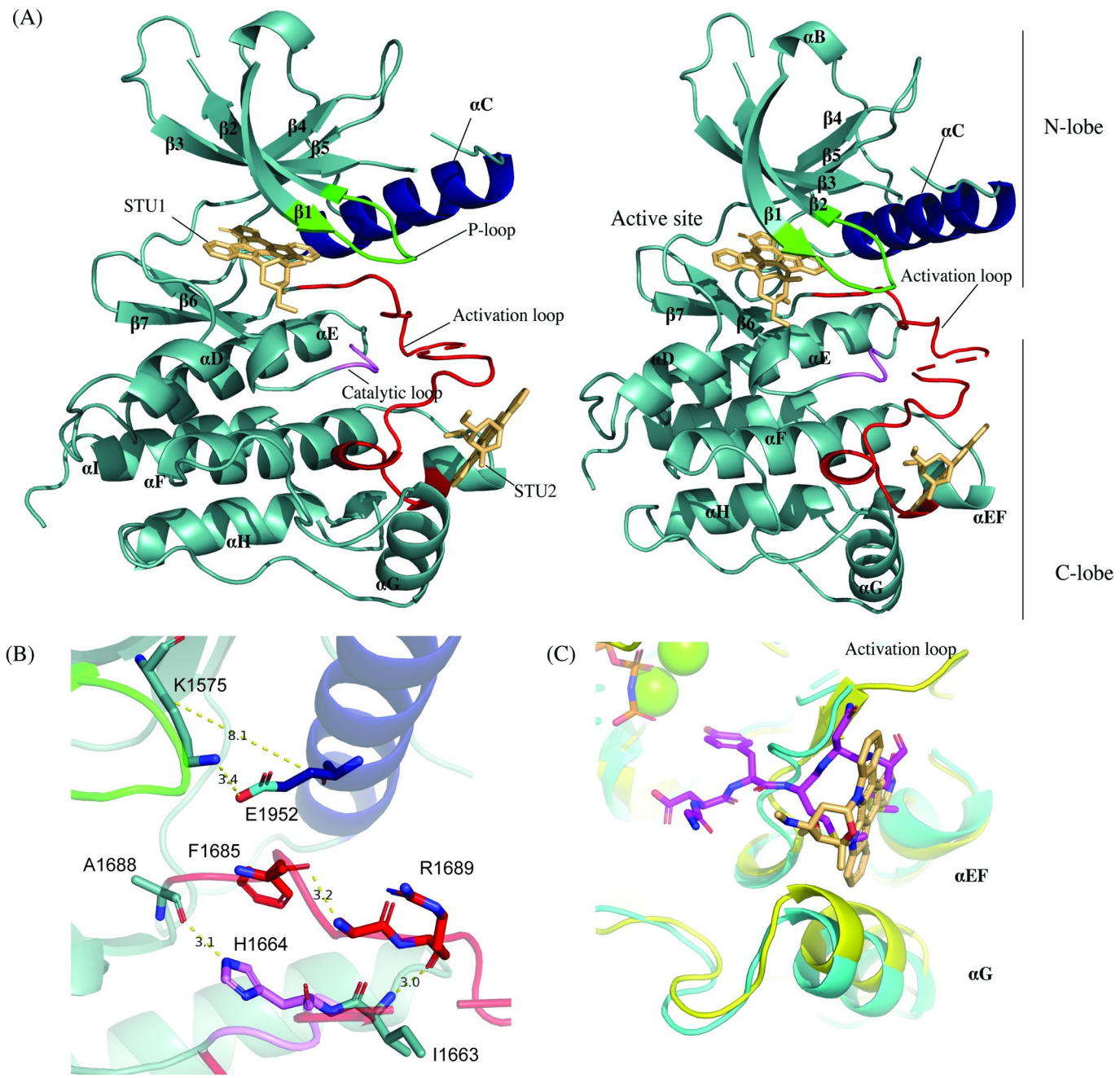

Fig 3. The RTKC8 kinase domain crystallized in an active conformation with staurosporine. (A) The RTKC8 structure reveals a classic, two lobed kinase domain (teal) with an αC helix (blue), P-loop (green), activation loop (red), catalytic loop (violet) with two staurosporines bound, one at the active site and the other at the peptide-substrate binding site (light orange). Two views are shown. The one on the right is rotated slightly from the left view so as to show the break in the activation loop more clearly. (**B**) The active αC-helix-in conformation is demonstrated by the presence of a salt bridge between the conserved αC-helix Glu1592 and Lys1575 in the β3 strand (distance between the two Cβs is <10Å; distance between the Lys ε-amnino and the Glu carboxyl is 3.4Å), the presence of a hydrogen bond between the backbone of the amino acid immediately prior to the DFG and the His in the HRD motif (3.1 Å) and an extended activation loop with a hydrogen bond between backbone carbonyl group of Arg1689 of the DFGxxR motif and the backbone amino group of Ile1663 which is immediately before the HRD motif. A hydrogen bond in the extended activation loop between the DFG motif Phe1685 backbone carbonyl and the amino nitrogen of Ala1688 (DFGxA, 3.2 Å) presents another feature of an active conformation of a kinase. (C) The second staurospaurine binds at the peptide-substrate binding site. When the structure of the human insulin kinase domain (yellow) in complex with a peptide substrate (1IR3) is overlayed on the RTKC8 kinase domain (cyan), the second staurospaurine (light orange) can be seen overlapping the position of the C-terminal segment of the substrate peptide (magenta) bound between the activation loop, the αEF helix and the αG helix.

https://doi.org/10.1371/journal.pone.0276413.g003

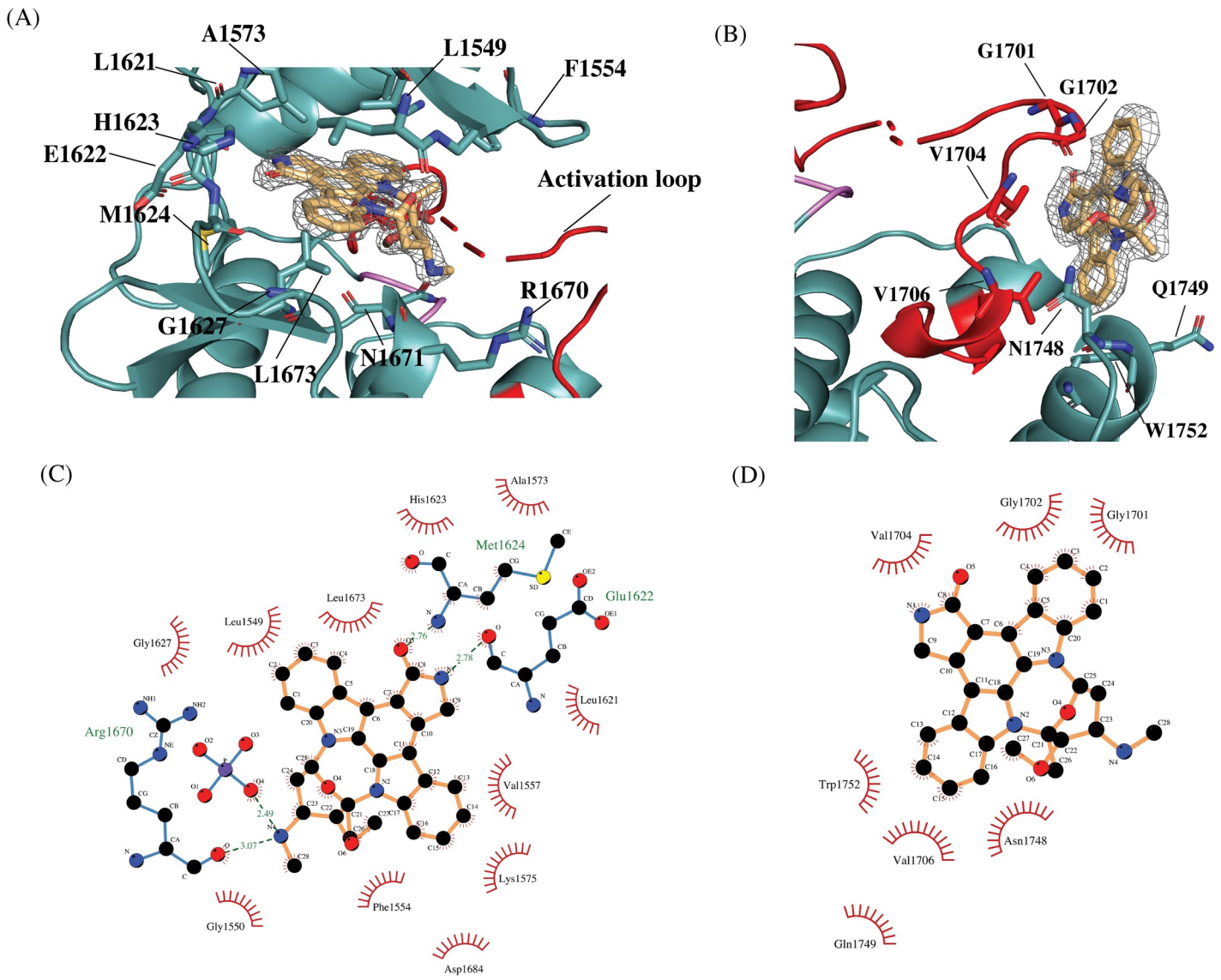

Fig 4. Two molecules of staurosporine (STU1 and STU2) are bound to the RTKC8 kinase. (A) One molecule of staurosporine (STU1) is bound at the active site, surrounded by the N-lobe β strands, the activation loop and the small helix between αE and β6 of the C-lobe (electron density for the Fo-Fc map with no compound modeled contoured at 2.0σ in grey). (B) The other staurosporine (STU2) is bound to the C-lobe, next to the activation loop, the αEF helix and the αG helix (electron density for the Fo-Fc map with no compound modeled contoured at 2.0σ in grey). (C) The contacts of STU1 in the active site of the RTKC8 kinase domain with side chain residues. Dotted lines (green) represent the hydrogen bonds and arches (red) represent the hydrophobic contacts (prepared by LigPlot+ [38]) (D) Contacts for the other saturosporine (STU2) prepared by LigPlot+.

https://doi.org/10.1371/journal.pone.0276413.g004

surrounded by β strands in the N-lobe, the glycine rich P-loop, the hinge region, the N-terminus of the activation loop and the small helix between αE and β6 of the C-lobe (Fig 4A). There are both hydrophobic and hydrophilic interactions between the inhibitor and residues inside the cleft. There are three hydrogen bonds, with Glu 1622, Met 1624 and Arg 1670 and five hydrophobic interactions, with Leu 1549, Val 1557 and Ala 1573 from N-lobe, and Gly 1627 and Leu 1673 from C-lobe. The hydrogens from the secondary amine group and the keto oxygen of the lactam ring of the inhibitor make a pair of hydrogen bonds with carbonyl oxygen on the backbone of Glu1622 and the primary amine on the backbone of Met1624 in the hinge region, respectively. The third hydrogen bond is observed between the methyloamino nitrogen

of staurosporine and the carbonyl oxygen on the backbone of Arg1670, in the ribose-binding pocket of the kinase domain (Fig 4B). This is a typical binding mode for staurospaurine in the active site of a kinase.

Another molecule of staurosporine was found bound to the C-lobe of the kinase domain in a hydrophobic groove between the activation loop, the αEF helix and the αG helix (Fig 4C). It also forms pi-pi stacking crystal contacts with its symmetry molecule, which implies that soaking of an apo or ATP/ADP bound crystal would not have revealed this secondary site, since the protein would likely have been different in crystals form without the inhibitor. This staurosporine is bound at the peptide-substrate binding site as revealed in the structure of the insulin receptor tyrosine kinase domain in complex with a peptide substrate (1IR3) [39] (Fig 3C). This site has been identified as a druggable site, [40] and is the site of a trans interaction in Aurora kinases, where the activation loop from one kinase binds to and presents as a substrate for phosphorylation by the other aurora kinase, so is designated the AAS site. This potential binding site has been identified in Src, Zap-70, and other tyrosine kinases using solvent mapping algorithms [40], and it is intriguing to see the inhibitor bound to a tyrosine kinas AAS site. The ligand interacts with hydrophobic residues (Gly 1701, Gly 1702, Val 1704, Val 1706) in the activation loop and Trp 1752, Asn 1748 and Gln 1749 in the αG-helix (Fig 4D).

Structural comparison of the RTKC8 kinase domain with the Ephrin kinase domain

A Dali search [41] of the RTKC8 kinase model (268 residues) from *Choanoflagellate* against the PDB database shows the highest structural similarity with the kinase domain of the Ephrin type A receptor 5 (2R2P, 295 residues) with the z-score of 36.2 and 38% sequence identity. The two structures were aligned over 214 atoms with a root mean square deviation of 0.84 Å (Fig 5). The Ephrin receptor juxtamembrane region modulates the kinase activity [42]. Two tyrosine residues in the juxtamembrane segment, when phosphorylated, release inhibition, and provide docking sites for SH2 domain-containing proteins to transmit the downstream signal [43]. However, the juxtamembrane domain of RTKC8 kinase does not contain tyrosine residues. In addition, while the kinase domains are structurally similar, the extracellular receptor on the Ephrin receptor bears no resemblance to that of the RTKC8 receptor. The structure of RTKC8 kinase demonstrates that the kinase domain remained highly conserved throughout the evolution of receptor tyrosine kinases from choanoflagellates to human but with divergent fused extracellular domains.

Concluding remarks

Choanoflagellates provide an excellent opportunity to study the evolution of receptor tyrosine kinases as they are the only group shown to have receptor tyrosine kinases outside of metazoans [9]. It is an enigma why such an apparently primitive organism has such an extensive kinome. We report the expression, purification, and initial characterization of the RTKC8 kinase domain from a choanoflagellate receptor tyrosine kinase. The isolated kinase domain is monomeric, which is usual for an isolated kinase domain, since it is usually the dimerization of the extracelluar sensor domains which dimerizes the kinase domains to activate them. We also present the crystal structure of the kinase domain in complex with staurospaurine. The purified RTKC8 kinase is catalytically active and can phosphorylate tyrosine residue containing peptides generated from its C-terminal tail sequence. This suggests that the tail tyrosine residues act as substrates for the kinase. These phosphorylated tyrosine residues most likely act as recruitment sites for a yet unrealized signaling pathway in choanoflagellates. What the role of these pathways in such a primitive organism is an intriguing question, open for further

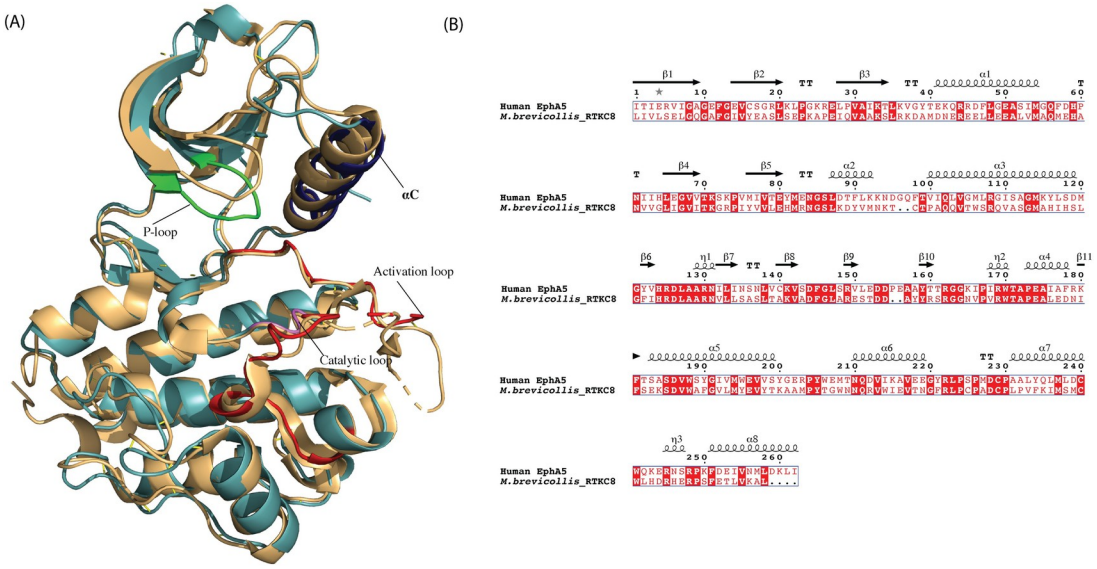

Fig 5. Structure and sequence alignment of RTKC8 kinase and EphA5 kinase. A Dali search identified the kinase domain of EphA5 (2R2P) as the structure with the highest similarity to RTKC8 in the PDB. The RTKC8 kinase (cyan) and EphA5 kinase (yellow) structures were aligned with a rmsd 0.836 Å. B. Sequence alignment of the EphA5 kinase domain with the RTKC8 kinase domain, with the secondary structural elements above the alignment. The alignment prepared with ESPript 3.0 [44].

https://doi.org/10.1371/journal.pone.0276413.g005

investigation. That Choanoflagellates have an Ephrin like kinase with a completely different sensor domain implies a divergence, in which either metazoans or choanoflagellates repurposed the kinase domain for their different signaling requirements. The unexpected secondary binding site for staurospaurine in the substrate binding site or Aurora activation segment (AAS), which is likely a non-optimal ligand for the site, shows the potential for this site in developing tyrosine kinase inhibitors.

Supporting information

S1 Fig. Cloning and Purification of Choan0RTKC8. (A). The expressed construct consists of an N-terminal hexa-His tag followed by a TEV protease site and the intracellular kinase domain flanked by residues from the juxtamembrane and tail (RTKC8 residues 1530–1811). (B) Chromatogram showing the protein eluted from gel filtration (Superdex 200 16/600 column) as a single symmetric peak at the expected elution volume for a monomer. (C) SDS-PAGE gel of fractions from gel filtration confirmed the protein purity.
(TIF)

S1 Raw images.
(TIF)

S1 File.
(PDF)

S2 File.
(MTZ)

S3 File.
(PDB)

Acknowledgments

We thank Xiaoxian Cao of the Kuriyan Lab for initial assistance with protein expression in insect cell culture. We thank Prof. Nicole King, UC Berkeley, for providing the Monosiga cDNA library. We thank the Berkeley Center for Structural Biology beamline staff at the Advanced Light Source, Lawrence Berkeley National Laboratory. We are thankful to Ruchika Bajaj and Neel H. Shah for providing comments and suggestions to improve the manuscript.

Author Contributions

Conceptualization: John Kuriyan.

Data curation: Teena Bajaj, Christine L. Gee.

Formal analysis: John Kuriyan, Christine L. Gee.

Funding acquisition: John Kuriyan.

Investigation: Teena Bajaj, John Kuriyan, Christine L. Gee.

Methodology: John Kuriyan, Christine L. Gee.

Project administration: John Kuriyan.

Resources: John Kuriyan.

Software: John Kuriyan.

Supervision: John Kuriyan, Christine L. Gee.

Validation: John Kuriyan, Christine L. Gee.

Visualization: Teena Bajaj, Christine L. Gee.

Writing – original draft: Teena Bajaj.

Writing – review & editing: John Kuriyan, Christine L. Gee.

References

1. Ullrich A, Schlessinger J. Signal transduction by receptors with tyrosine kinase activity. Cell. 1990; 61: 203–212. https://doi.org/10.1016/0092-8674(90)90801-k PMID: 2158859
2. Hubbard SR. Structural analysis of receptor tyrosine kinases. Prog Biophys Mol Biol. 1999; 71: 343–358. https://doi.org/10.1016/s0079-6107(98)00047-9 PMID: 10354703
3. Pawson T, Scott JD. Signaling through scaffold, anchoring, and adaptor proteins. Science. 1997; 278: 2075–2080. https://doi.org/10.1126/science.278.5346.2075 PMID: 9405336
4. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010; 141: 1117–1134. https://doi.org/10.1016/j.cell.2010.06.011 PMID: 20602996
5. Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2000; 103: 211–225. https://doi.org/10.1016/s0092-8674(00)00114-8 PMID: 11057895
6. King N. Choanoflagellates. Curr Biol. 2005; 15: R113–4. https://doi.org/10.1016/j.cub.2005.02.004 PMID: 15723775
7. King N. The unicellular ancestry of animal development. Dev Cell. 2004; 7: 313–325. https://doi.org/10.1016/j.devcel.2004.08.010 PMID: 15363407
8. Fairclough S, King N. Choanoflagellates. Choanoflagellida, collared-flagellates. In: The Tree 383 of Life Web Project [Internet]. 2006 [cited 8 Mar 2022]. Available: http://tolweb.org/Choanoflagellates/2375
9. King N, Carroll SB. A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. Proc Natl Acad Sci USA. 2001; 98: 15032–15037. https://doi.org/10.1073/pnas.261477698 PMID: 11752452
10. King N, Hittinger CT, Carroll SB. Evolution of key cell signaling and adhesion protein families predates animal origins. Science. 2003; 301: 361–363. https://doi.org/10.1126/science.1083853 PMID: 12869759

11. Manning G, Young SL, Miller WT, Zhai Y. The protist, Monosiga brevicollis, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. Proc Natl Acad Sci USA. 2008; 105: 9674–9679. https://doi.org/10.1073/pnas.0801314105 PMID: 18621719

12. Yeung W, Kwon A, Taujale R, Bunn C, Venkat A, Kannan N. Evolution of functional diversity in the holozoan tyrosine kinome. Mol Biol Evol. 2021; 38: 5625–5639. https://doi.org/10.1093/molbev/msab272 PMID: 34515793

13. Callebaut I, Gilgès D, Vigon I, Mornon JP. HYR, an extracellular module involved in cellular adhesion and related to the immunoglobulin-like fold. Protein Sci. 2000; 9: 1382–1390. https://doi.org/10.1110/ps.9.7.1382 PMID: 10933504

14. Wang Q, Crevenna AH, Kunze I, Mizuno N. Structural basis for the extended CAP-Gly domains of p150 (glued) binding to microtubules and the implication for tubulin dynamics. Proc Natl Acad Sci USA. 2014; 111: 11347–11352. https://doi.org/10.1073/pnas.1403135111 PMID: 25059720

15. Barker SC, Kassel DB, Weigl D, Huang X, Luther MA, Knight WB. Characterization of pp60c-src tyrosine kinase activities using a continuous assay: autoactivation of the enzyme is an intermolecular autophosphorylation process. Biochemistry. 1995; 34: 14843–14851. https://doi.org/10.1021/bi00045a027 PMID: 7578094

16. Shah NH, Wang Q, Yan Q, Karandur D, Kadlecek TA, Fallahee IR, et al. An electrostatic selection mechanism controls sequential kinase signaling downstream of the T cell receptor. Elife. 2016; 5: e20105. https://doi.org/10.7554/eLife.20105 PMID: 27700984

17. Shah NH, Löbel M, Weiss A, Kuriyan J. Fine-tuning of substrate preferences of the Src-family kinase Lck revealed through a high-throughput specificity screen. Elife. 2018; 7: e35190. https://doi.org/10.7554/eLife.35190 PMID: 29547119

18. Newman J, Egan D, Walter TS, Meged R, Berry I, Ben Jelloul M, et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. Acta Crystallogr Sect D, Biol Crystallogr. 2005; 61: 1426–1431. https://doi.org/10.1107/S0907444905024984 PMID: 16204897

19. Kabsch W. XDS. Acta Crystallogr Sect D, Biol Crystallogr. 2010; 66: 125–132. https://doi.org/10.1107/S0907444909047337 PMID: 20124692

20. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr Sect D, Biol Crystallogr. 2010; 66: 133–144. https://doi.org/10.1107/S0907444909047374 PMID: 20124693

21. Evans P. Scaling and assessment of data quality. Acta Crystallogr Sect D, Biol Crystallogr. 2006; 62: 72–82. https://doi.org/10.1107/S0907444905036693 PMID: 16369096

22. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr Sect D, Biol Crystallogr. 2013; 69: 1204–1214. https://doi.org/10.1107/S0907444913000061 PMID: 23793146

23. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D, Biol Crystallogr. 2011; 67: 235–242. https://doi.org/10.1107/S0907444910045749 PMID: 21460441

24. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr Sect D, Biol Crystallogr. 2007; 63: 32–41. https://doi.org/10.1107/S0907444906045975 PMID: 17164524

25. Wilson C, Agafonov RV, Hoemberger M, Kutter S, Zorba A, Halpin J, et al. Kinase dynamics. Using ancient protein kinases to unravel a modern cancer drug's mechanism. Science. 2015; 347: 882–886. https://doi.org/10.1126/science.aaa1823 PMID: 25700521

26. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Sect D, Biol Crystallogr. 2010; 66: 213–221. https://doi.org/10.1107/S0907444909052925 PMID: 20124702

27. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr Sect D, Biol Crystallogr. 2010; 66: 486–501. https://doi.org/10.1107/S0907444910007493 PMID: 20383002

28. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 2018; 27: 293–315. https://doi.org/10.1002/pro.3330 PMID: 29067766

29. Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, et al. Collaboration gets the most out of software. Elife. 2013; 2: e01456. https://doi.org/10.7554/eLife.01456 PMID: 24040512

30. McDonald NQ, Murray-Rust J, Blundell TL. The first structure of a receptor tyrosine kinase domain:a further step in understanding the molecular basis of insulin action. Structure. 1995; 3: 1–6. https://doi.org/10.1016/S0969-2126(01)00129-0 PMID: 7743124

31. Fabbro D, Cowan-Jacob SW, Moebitz H. Ten things you should know about protein kinases: IUPHAR Review 14. Br J Pharmacol. 2015; 172: 2675–2700. https://doi.org/10.1111/bph.13096 PMID: 25630872

32. Yamaguchi H, Hendrickson WA. Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. Nature. 1996; 384: 484–489. https://doi.org/10.1038/384484a0 PMID: 8945479

33. Modi V, Dunbrack RL. Defining a new nomenclature for the structures of active and inactive kinases. Proc Natl Acad Sci USA. 2019; 116: 6818–6827. https://doi.org/10.1073/pnas.1814279116 PMID: 30867294

34. Huse M, Kuriyan J. The conformational plasticity of protein kinases. Cell. 2002; 109: 275–282. https://doi.org/10.1016/s0092-8674(02)00741-9 PMID: 12015977

35. Pucheta-Martínez E, Saladino G, Morando MA, Martinez-Torrecuadrada J, Lelli M, Sutto L, et al. An Allosteric Cross-Talk Between the Activation Loop and the ATP Binding Site Regulates the Activation of Src Kinase. Sci Rep. 2016; 6: 24235. https://doi.org/10.1038/srep24235 PMID: 27063862

36. Lawrie AM, Noble ME, Tunnah P, Brown NR, Johnson LN, Endicott JA. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. Nat Struct Biol. 1997; 4: 796–801. https://doi.org/10.1038/nsb1097-796 PMID: 9334743

37. Jin L, Pluskey S, Petrella EC, Cantin SM, Gorga JC, Rynkiewicz MJ, et al. The three-dimensional structure of the ZAP-70 kinase domain in complex with staurosporine: implications for the design of selective inhibitors. J Biol Chem. 2004; 279: 42818–42825. https://doi.org/10.1074/jbc.M407096200 PMID: 15292186

38. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995; 8: 127–134. https://doi.org/10.1093/protein/8.2.127 PMID: 7630882

39. Hubbard SR. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J. 1997; 16: 5572–5581. https://doi.org/10.1093/emboj/16.18.5572 PMID: 9312016

40. Yueh C, Rettenmaier J, Xia B, Hall DR, Alekseenko A, Porter KA, et al. Kinase atlas: druggability analysis of potential allosteric sites in kinases. J Med Chem. 2019; 62: 6512–6524. https://doi.org/10.1021/acs.jmedchem.9b00089 PMID: 31274316

41. Holm L, Laakso LM. Dali server update. Nucleic Acids Res. 2016; 44: W351–5. https://doi.org/10.1093/nar/gkw357 PMID: 27131377

42. Davis TL, Walker JR, Loppnau P, Butler-Cole C, Allali-Hassani A, Dhe-Paganon S. Autoregulation by the juxtamembrane region of the human ephrin receptor tyrosine kinase A3 (EphA3). Structure. 2008; 16: 873–884. https://doi.org/10.1016/j.str.2008.03.008 PMID: 18547520

43. Liang L-Y, Patel O, Janes PW, Murphy JM, Lucet IS. Eph receptor signalling: from catalytic to non-catalytic functions. Oncogene. 2019; 38: 6567–6584. https://doi.org/10.1038/s41388-019-0931-2 PMID: 31406248

44. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014; 42: W320–W324. https://doi.org/10.1093/nar/gku316 PMID: 24753421