Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

Research Article

N-terminal phosphorylation regulates the activity of glycogen synthase kinase 3 from *Plasmodium falciparum*

Samuel Pazicky[1,2], Arne Alder[1,3,4], Haydyn Mertens[2], Dmitri Svergun[2], Tim Gilberger[1,3,4] and Christian Löw[1,2]

[1]Centre for Structural Systems Biology (CSSB), Notkestrasse 85, D-22607 Hamburg, Germany; [2]Molecular Biology Laboratory (EMBL), Hamburg Unit c/o Deutsches Elektronen Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany; [3]Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Strasse 74, D-20359 Hamburg, Germany; [4]Department of Biology, University of Hamburg, Hamburg, Germany

Correspondence: Christian Löw (christian.loew@embl-hamburg.de)

As the decline of malaria cases stalled over the last five years, novel targets in *Plasmodium falciparum* are necessary for the development of new drugs. Glycogen Synthase Kinase (PfGSK3) has been identified as a potential target, since its selective inhibitors were shown to disrupt the parasitès life cycle. In the uncanonical N-terminal region of the parasite enzyme, we identified several autophosphorylation sites and probed their role in activity regulation of PfGSK3. By combining molecular modeling with experimental small-angle X-ray scattering data, we show that increased PfGSK3 activity is promoted by conformational changes in the PfGSK3 N-terminus, triggered by N-terminal phosphorylation. Our work provides novel insights into the structure and regulation of the malarial PfGSK3.

Introduction

Plasmodium falciparum is an intracellular eukaryotic parasite that causes the most severe form of malaria in humans, infecting over 200 million people every year [1]. The emergence of malaria parasites resistant to all currently used antimalarial drugs imposes a serious threat to public health [2]. Identification of new drug targets and drugs is therefore crucial for future handling of the malaria pandemics. Kinases comprise up to 10% of all drug targets [3] and a number of potent inhibitors were discovered that are directed against *P. falciparum* kinases, such as glycogen synthase kinase 3 (PfGSK3) [4]. Human GSK3ß plays a crucial role in a number of signaling processes, regulating glycogen metabolism [5–7], cell cycle and growth [8–10], translation [7], embryonic development [9,11] or differentiation of neurons [12–16]. Consequently, the inhibitors of GSK3 are investigated as potential treatment options for neurodegenerative and psychiatric diseases [17–23] and its structure, catalytic mechanism and regulation are well studied [24–26]. GSK3β is thought to be constitutively activated by autophosphorylation at Y216 [27–29] and inactivated by phosphorylation of N-terminal S9 mediated by several protein kinases, such as PKA and PKB [6,30]. The inactivation is achieved as the phosphorylated S9 enables binding of GSK3 N-terminus to a pocket for a priming phosphorylation in the substrate-binding region and thus reduces the binding affinity for substrates [24,25].

In *P. falciparum*, two genes encode GSK3 (PF3D7_0312400 and PF3D7_1316000) and these enzymes are much less understood in their physiological role and in the molecular mechanism. GSK3ß (PF3D7_0312400, further as PfGSK3) is the only isoform that is expressed in blood stages [31] of the parasite and is involved in the regulation of red blood cell invasion [32]. This essential process is parasite-driven and relies on the orchestrated interplay of multiple protein–protein interactions [33,34]. PfGSK3 is involved in activation of apical membrane antigen 1 (AMA1) that subsequently mediates the formation of the so called 'tight junction' with the host cell prior to the invasion [32,35–40]. In this two-step phosphorylation event, protein kinase A (PfPKA) first phosphorylates

Received: 7 December 2021
Revised: 9 January 2022
Accepted: 12 January 2022

Accepted Manuscript online: 13 January 2022
Version of Record published: 4 February 2022

Biochemical Journal (2022) 479 337–356
https://doi.org/10.1042/BCJ20210829

S610 of AMA1 [41–43], which in turn enables subsequent phosphorylation of AMA1 residue T613 by PfGSK3 [32]. Thus, PfGSK3 serves as a secondary kinase, although both phosphorylation sites are equally important for AMA1 function and *P. falciparum* invasion [32,36,40].

PfGSK3 inhibitors, built on a thieno[2,3-b]pyridine or benzofuran scaffold, have been developed that display selective antiplasmodial activity with IC_{50} values in the low micromolar to sub-micromolar range [44–46] and one of these inhibitors has been shown to abrogate red blood cell invasion [32]. Further improvements of these inhibitors could be guided by a three-dimensional structure of PfGSK3 that has, however, not been determined so far. Based on sequence homology across GSK3 proteins (Supplementary Figure S1A), PfGSK3 consists of two conserved structural domains: a β-sheet-rich lobe with an ATP binding pocket and an α-helical lobe with the activation loop and a substrate binding site. These domains are typically preceded by an N-terminal domain that is predicted to be disordered [25]. In PfGSK3, the important residues in the ATP binding pocket and in the activation loop remain conserved, however, its N-terminus displays an unique extension and its function remains unknown [47].

To analyze this in more detail, we performed thorough structural and functional characterization of PfGSK3 *in vitro*. We provide evidence that PfGSK3 exhibits autophosphorylation of the uncanonical N-terminal domain. We show that phosphorylation of the N-terminal residues induces conformational changes of the protein that lead to an increase in the enzymatic activity of PfGSK3. Furthermore, we reveal that binding of heavy metal ions such as zinc cations to PfGSK3 induce a reversible formation of large complexes that result in a complete loss of its enzymatic activity.

Methods

Cloning and mutagenesis

The vector with N-terminally His-tagged PfGSK3 was generated by PCR amplification of the GSK3 coding sequence from *P. falciparum* cDNA followed by Ligation Independent Cloning into HindIII/KpnI-cleaved plasmid pOPIN F [48] using the In-Fusion HD EcoDry Cloning Kit (Takara Clontech) according to the manufacturer's instructions. The mutants S226A, Y229A and S226A/Y229A were generated by overlap extension PCR amplification from the original vector and Ligation Independent Cloning as described above. The wild-type protein and the mutant K96A cloned in pET28a vector were ordered from GenScript. The N-terminally truncated constructs were cloned by amplifying the sequence from the original vector and subcloning into BsaI-cleaved plasmid pNIC28_Bsa4 by SLiCE cloning [49].

PfGSK3 expression and purification

E. coli C41(DE3) culture transformed with the PfGSK3 coding vector was grown in TB medium supplemented with 0.1 mg/ml ampicillin at 37°C, induced at $OD_{600nm} = 0.7$ with 0.5 mM IPTG and harvested after 4 h. The pellets were resuspended in 5 ml of lysis buffer (20 mM NaP pH 7.5, 300 mM NaCl, 15 mM imidazole, 5% glycerol, 0.5 mM TCEP, 1 mg/ml lysozyme, 5 U/ml DNase, 1 Roche protease inhibitor tablet/100 ml) per 1 g of wet weight culture. The suspension was passed three times through an emulsifier at 15 000 psi, centrifuged at 19 000***g*** and the supernatant was loaded on a pre-equilibrated His-Trap column. The column was washed with lysis buffer supplemented with 40 mM imidazole and the protein was eluted at increasing imidazole concentrations. The protein containing fractions were concentrated (10 kDa cut-off concentrator) with the addition of 1 mg of 3C protease per 3 L of culture volume and further separated on a Superdex 200 HiLoad column (GE Healthcare) in 50 mM Tris pH 8.0, 20 mM NaCl, 0.5 mM TCEP. The protein containing fractions were concentrated again, loaded on ResourceQ anion exchange column (GE Healthcare) and eluted with increasing concentration of NaCl. Fractions representing different PfGSK3 species were pooled together, concentrated, dialyzed against the analysis buffer (typically 20 mM Tris pH 8.0, 100 mM NaCl, 0.5 mM TCEP) and flash-frozen in liquid nitrogen and stored at −80°C. All purification steps were performed at 4°C.

Expression and purification of insoluble PfGSK3 constructs

The insoluble PfGSK3 constructs (K96A, S226A/Y229A, ΔN64) were expressed as described for the wild-type PfGSK3. The pellets of harvested bacteria were resuspended in 5 ml of lysis buffer (20 mM NaP pH 7.5, 300 mM NaCl, 15 mM imidazole, 5% glycerol, 0.5 mM TCEP, 1 mg/ml lysozyme, 5 U/ml DNase, 1 Roche protease inhibitor tablet/100 ml) per 1 g of wet weight culture. The suspension was passed three times through an emulsifier at 15 000 psi, centrifuged at 19 000***g*** and the pellet was dissolved in a denaturating buffer (20 mM

NaP (pH 7.5), 300 mM NaCl, 25 mM imidazole, 5% glycerol, 6 M urea, 0.5 mM TCEP) and incubated with 1 ml NiNTA bead slurry (Sigma) for 1 h, 4°C. The beads were subsequently washed twice with 10 ml of the same buffer and subsequently eluted with 5 ml of the same buffer containing extra 250 mM imidazole. The protein was concentrated (10 kDa c/o) and the sample for mass spectrometry was separated by SDS–PAGE gel.

Small-scale expression tests

An amount of 4 ml of TB supplemented with ampicillin was inoculated wit pre-cultured *E. coli* C41(DE3) that had been transformed with the tested plasmids. The cultures were grown at 37°C until the $OD_{600nm} = 0.6$ and further either induced with 0.5 mM IPTG and grown for 4 h at 37°C or 25°C, or induced with 0.1 mM IPTG and grown for 16 h at 18°C. $OD_{600\ nm}$ was measured and culture volume corresponding to $OD_{600\ nm}$ of 2.0/ml was harvested. The pellets were resuspended in 400 μl of lysis buffer (lysate fractions), then lysed by 10 min of vortexing with 100 μl glass beads and the cell debris were removed by centrifugation (soluble fractions). An amount of 5 μl of each fraction was mixed with 10 μl of loading dye and run on SDS–PAGE in duplicates. One gel of each set of samples was blotted on PVDF membrane (*Bio-Rad*). The membranes were blocked by 1% BSA in TBS-T, washed, incubated for 1 h with a HisProbe-HRP conjugated antibody (ThermoFisher), washed with TBS-T and the blot was developed using Super Signal West Pico Substrate (ThermoFisher) and Super Signal West Femto Substrate (ThermoFisher) in a 1 : 10 ratio.

Glutaraldehyde cross-linking

An amount of 1 mg/ml PfGSK3 in PBS and 0.5 mM TCEP was mixed with 8× molar excess of either EDTA or $ZnCl_2$ and 0.01% glutaraldehyde. The reaction was quenched with 330 mM Tris (final concentration) and SDS–PAGE loading dye and SDS–PAGE was run for 150 min at 150 V. The gels were stained with Direct Blue stain.

Thermal shift assay

The stability of PfGSK3 under different buffer conditions and additives was measured by nanoDSF (Prometheus NT.48, NanoTemper Technologies, GmbH). The protein concentration was adjusted to 1 mg/ml. 10 μl of samples were loaded in the glass capillaries and heated from 20°C to 90°C with a heating rate of 1°C/min. The fluorescence signals with excitation wavelength of 280 nm and emission wavelengths of 330 and 350 nm were recorded and the melting temperature was calculated as maximum of the derivative of the fluorescence at 330 nm.

Circular dichroism

Circular dichroism was measured on a Chirascan CD spectrometer (Applied Photophysics). The protein concentration was adjusted to 1 μM by 2 mM NaP (pH 7.5), 4 mM NaCl, 0.05 mM TCEP prior to the measurement. The circular dichroism spectrum was measured 10× between 185 nm and 260 nm with 1 nm step in 1 mm quartz cuvette and analyzed by the CDSSTR algorithm [50,51] using DichroWeb [52].

Analytical size exclusion chromatography

The analytical size exclusion chromatography was performed on an Agilent Bio-LC system using a home-packed Superose 5/150 column (25 μl sample). GSK3 was mixed with zinc chloride to achieve molar ratios ranging from 1 : 8 to 8 : 1 at fixed final GSK3 concentration of 1 mg/ml. PBS supplemented with 0.5 mM TCEP was used as a mobile phase and the resulting chromatograms were recorded using the as fluorescence signal with excitation wavelength of 280 nm and emission wavelength of 350 nm as readout.

Autophosphorylation IEX assay

An amount of 0.5 mg/ml of PfGSK3 was incubated at 37°C with 2 mM ATP or cAMP and 10 mM $MgCl_2$ overnight. The samples were subsequently dialyzed at 4°C overnight in 50 mM Tris pH 8.0, 20 mM NaCl, 0.5 mM TCEP. The samples were analyzed by analytical ion exchange chromatography using a Resource Q column (GE Healthcare) and the 1260 Infinity Bio-inert high-performance liquid chromatography system (Agilent Technologies) at 10°C. The system and column were equilibrated in 50 mM Tris pH 8.0, 20 mM NaCl, 0.5 mM TCEP. 100 μl of sample was injected and eluted with an increasing concentration of NaCl. The system was run at 0.2 ml/min ad the elution profile was analyzed by UV fluorescence detector with absorbance at 280 nm and emission at 350 nm.

Biochemical Journal (2022) 479 337–356
https://doi.org/10.1042/BCJ20210829

Autophosphorylation gel assay

The protein concentration of separated IEX peak fractions was adjusted to 0.25 mg/ml and 20 μl of each fraction was mixed with 10 μl of 3× loading dye. An amount of 10 μl of samples were run on SDS–PAGE next to each other twice. One gel was stained with Direct Blue stain (C.B.S. Scientific) and the other gel was stained with Pro-Q Diamond stain (Thermofisher). The stained gels were imaged with Imager Gel Doc XR+ System (Bio-Rad) and the band intensity estimated with Image Lab software (Bio-Rad). The relative intensity of the bands was first estimated for each gel separately (relative to the band in lane 1) and the relative phosphorylation was calculated as the ratio between relative intensity of Direct Blue-stained bands and relative intensity of Pro-Q Diamond-stained bands.

Dynamic light scattering

Dynamic light scattering was measured 30× on Wyatt Dynapro Plate Reader II with 1 mg/ml GSK3 in 20 mM Tris (pH 8.0), 100 mM NaCl, 0.5 mM TCEP. The molar mass was assessed from hydrodynamic radius (calculated from cumulants fit from autocorrelation functions averaged over 30 measurements), particle density and globular particle conformation model in DYNAMICS software (Wyatt Technology).

Small angle X-ray scattering

The SEC-SAXS data were collected at the P12 BioSAXS Beamline at the PETRA III storage ring (c/o DESY, Hamburg, Germany) [53]. The concentrated samples were dialyzed against 20 mM Tris pH 8.0, 100 mM NaCl, 0.5 mM TCEP overnight at 4°C and the concentration was adjusted to 10 mg/ml. The X-ray scattering data were measured in an on-line SEC-SAXS mode, using a SD200 Increase 5/150 column (GE Healthcare) at 0.4 ml/min with 1 frame recorded per second (Table 1). The data were automatically processed using the SASFLOW pipeline [54] and further analyzed using the ATSAS suite [55] of programs to determine the molar mass, radius of gyration and real-space distance distribution. Scattering curves were computed from models using CRYSOL [56] with a constant subtraction and maximum order of harmonics set to 50. The rigid body modeling performed using CORAL [57], setting different parts of models yielding best CRYSOL fits as disordered. These were residues 1–65 and 403–440 in conformation A, resides 47–65 and 403–440 in conformation B and residues 403–440 in conformation C. The raw data were uploaded to SASBDB under accession numbers SASDL77 and SASDL87.

Kinase assay

To investigate enzymatic activity of recombinant PfGSK3 a commercial luminescence-based kinase assay (KinaseGlo Plus, Promega) was used as previously described [58]. Briefly, 20 ng PfGSK3, 12 μM GS-1 peptide substrate (residues 636–661 of the human glycogen synthase 1; sequence YRRAAVPPSPSLSRHSSPHQ(pS)EDEEE; pS = prephosphorylated serine, Promega) and 6 μM ATP (UltraPure, Promega) in the kinase reaction buffer (40 mM Tris–HCl pH 7.5; 20 mM $MgCl_2$; 0.1 mg/ml BSA) were used as standard reaction mix (final reaction volume 5 μl). The kinase reaction was performed for 30 min at 30°C followed by incubation with KinaseGlo reagent according to the manufacturer's instructions. The reaction was transferred to a solid white 384-well plate (NUNC, ThermoFisher) and the luminescence signal was measured in an EnVision Multilabel Plate Reader (PerkinElmer, Integration time 0.5 sec/well). To investigate reaction kinetics, kinase and substrate concentrations were varied between 0 and 100 ng or 0 and 15 μM, respectively. To investigate autophosphorylation, reactions were performed with kinase concentrations between 0 and 500 ng, 30 μM ATP and without a substrate.

Sample preparation for mass spectrometry

To assess the phosphorylation status of the recombinant PfGSK3 (purified protein, separated IEX fractions, insoluble constructs), each sample was separated on SDS–PAGE gel. Coomassie-stained bands were excised, cut into small pieces and transferred to 0.5 ml Eppendorf tubes. For all following steps, buffers were exchanged by two consecutive 15 min incubation steps of the gel pieces with 200 μl of acetonitrile (ACN), whereby the ACN was removed after each step. Proteins were reduced by the addition of 200 μl of a 10 mM DTT solution in 100 mM ammonium bicarbonate (AmBiC, Sigma–Aldrich, A6141) and incubation at 56°C for 30 min. Proteins were alkylated by the addition of 200 μl of 55 mM chloroacetamide (CAA), 100 mM AmBiC solution and incubation for 20 min in the dark. A 0.1 μg/μl stock solution of trypsin (Promega, V511A) in trypsin

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

Table 1 SAXS sample details, data acquisition parameters, structural parameters and atomistic modeling

Sample details		
Sample	PfGSK3 — F1	PfGSK3 — F4
Organism	*P. falciparum*	
Source	*E. coli* (DE3) C41	
UniProt ID	O77344	
Extinction coefficient ε (at 280 nm, M^{-1} cm^{-1})	48 250	
Molecular weight from chemical composition (Da)	52 045.2	
Concentration (analysis or injection, mg/ml)	10	
Solvent composition	20 mM Tris pH 8.0, 100 mM NaCl, 0.5 mM TCEP	
SAS data collection parameters		
Beamline	P12, DESY/EMBL, Hamburg (Germany)	
Detector	Pilatus 6M	
Energy (kEV)	10.0	
Sample-to-detector distance (mm)	3000	
q-measurement range ($Å^{-10}$)	0.003–0.732	
Absolute scaling method	Relative to the scattering of pure water	
Method for monitoring radiation damage	Frame comparison	
Exposed time for frame	900 s (900 × 1 s)	900 s (900 × 1 s)
Mode	SEC-SAXS	SEC-SAXS
Sample temperature (°C)	20	20
Structural parameters		
Guinier analysis		
I(0) (cm^{-1})	0.035 ± 0.001	0.021 ± 0.001
R_g (Å)	32.7 ± 0.3	31.5 ± 0.2
q-range ($Å^{-1}$)	0.01–0.04	0.01–0.04
Fidelity (Quality of fit parameter, *AutoRg*)	0.69	0.98
Molecular weight (kDa)[1]		
From V_c	64.0	48.7
From *MoW*	77.2	34.1
P(r) analysis		
I(0) (cm^{-1})	0.035 ± 0.001	0.021 ± 0.001
R_g (Å)	3.9 ± 0.01	32.5 ± 0.01
D_{max} (Å)	130 ± 10	115 ± 10
q-range ($Å^{-1}$)	0.01–0.50	0.01–0.25
Total quality estimate	0.69	0.82
Porod volume × 10^3 ($Å^3$)	102	101
Atomistic modeling[2]	*CRYSOL* with constant subtraction and maximum order of harmonics equal to 50	*CORAL* hybrid rigid body modeling
Structures	Robetta model WLT	
q-range for modeling ($Å^{-1}$)	0.01–0.50	0.01–0.58
χ^2	1.23	1.06
Predicted R_g (Å)	33.8	
Starting crystal structures		Robetta model 3xS/E
Flexible residues		47–63, 403–440
SASBDB IDs for data and models		
	SASDMC9	SASDL87

[1]Calculated using the appropriate functions in ATSAS [55];
[2]The values for the best model for each respective dataset are shown.

Biochemical Journal (2022) 479 337–356
https://doi.org/10.1042/BCJ20210829

resuspension buffer (Promega, V542A) was diluted with ice-cold 50 mM AmBiC buffer to achieve a final concentration of 1 ng/μl. 50 μl thereof were added to gel pieces, which were incubated first for 30 min on ice and then over night at 37°C. Gel pieces were sonicated for 15 min, spun down and the supernatant was transferred into a glass vial (VDS Optilab, 93908556). Remaining gel pieces were washed with 50 μl of an aqueous solution of 50% ACN and 1% formic acid and sonicated for 15 min. The combined supernatants were dried in a speed-vac and reconstituted in 10 μl of an aqueous solution of 0.1% (v/v) formic acid.

Mass spectrometry data acquisition

The mass spectra of the initial phosphorylation site assessment of PfGSK3 (Supplementary Figure S4B) were acquired as follows: An UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (Precolumn C18 PepMap100, 5 mm, 300 μm i.d., 5 μm, 100 Å) and an analytical column (Acclaim PepMap 100. 75 × 50 cm C18, 3 mm, 100 Å) was used. Trapping was carried out with a constant flow of trapping solution (0.05% trifluoroacetic acid in water) at 30 μl/min onto the trapping column for 6 min. Subsequently, peptides were eluted via the analytical column running solvent A (0.1% formic acid in water) with a constant flow of 0.3 μl/min, with increasing percentage of solvent B (0.1% formic acid in acetonitrile) from 2% to 4% in 4 min, from 4% to 8% in 2 min, then 8% to 28% for a further 37 min, in another 9 min. from 28%–40%, and finally 40%–80% for 3 min followed by re-equilibration back to 2% B in 5 min. The outlet of the analytical column was coupled directly to an Orbitrap QExactive™ plus Mass Spectrometer (Thermo) using the Nanospray Flex™ ion source in positive ion mode. The QExactive plus was operated in positive ion mode with a spray voltage of 2.2 kV and capillary temperature of 275°C. Full scan MS spectra with a mass range of 350–1400 m/z were acquired in profile mode using a resolution of 70 000 (maximum injections time of 100 ms, AGC target was set to 3×10^6 ions). Precursors were isolated using the quadrupole with a window of 1.7 m/z and fragmentation was triggered by HCD in fixed collision energy mode with fixed collision energy of 28%. MS2 spectra were acquired with the Orbitrap with a resolution of 17.500 and a max injection time of 70 ms.

The mass spectra of all other experiments were recorded as follows: Peptides were analyzed by LC–MS/MS on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scentific) as previously described [59]. To this end, peptides were separated using an Ultimate 3000 nano RSLC system (Dionex) equipped with a trapping cartridge (Precolumn C18 PepMap100, 5 mm, 300 μm i.d., 5 μm, 100 Å) and an analytical column (Acclaim PepMap 100. 75 × 50 cm C18, 3 mm, 100 Å) connected to a nanospray-Flex ion source. For the detection of posttranslational modified peptides, peptides were loaded onto the trap column at 30 μl per min using solvent A (0.1% formic acid) and peptides were eluted using a gradient from 2 to 80% Solvent B (0.1% formic acid in acetonitrile) over 60 min at 0.3 μl per min (all solvents were of LC–MS grade). The Orbitrap Fusion Lumos was operated in positive ion mode with a spray voltage of 2.4 kV and capillary temperature of 275°C. Full scan MS spectra with a mass range of 375–1200 m/z were acquired in profile mode using a resolution of 120 000 (maximum injections time of 50 ms, AGC target was set to Standard with a max injection time of 50 ms. Precursors were isolated using the quadrupole with a window of 1.2 m/z and fragmentation was triggered by HCD in fixed collision energy mode with fixed collision energy of 34%. MS2 spectra were acquired with the Orbitrap with a resolution of 15.000 and a max injection time of 54 ms.

Mass spectrometry data analysis

Acquired data were analyzed using IsobarQuant [60] and Mascot V2.4 (Matrix Science) using a reverse UniProt FASTA database from *E. coli* (UP000000625) or *Plasmodium falciparum* (UP000001450) including common contaminants. The following modifications were taken into account: Carbamidomethyl (C, fixed), Acetyl (K, variable), Acetyl (Protein N-term, variable), Oxidation (M, variable), as well as the variable modification Phospho (ST) and Phospho (Y). The mass error tolerance for full scan MS spectra was set to 10 ppm and to 0.02 Da for MS/MS spectra. A maximum of two missed cleavages were allowed. A minimum of two unique peptides with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were required on the peptide and protein level [61]. Prior to performing the search, the protein sequences corresponding to the expressed PfGSK3 constructs were added to the database for the expression host. The raw data have been uploaded to the PRIDE database with the dataset identifier PXD026387.

Correlation analysis

Each biological replicate of mass spectrometry experiment was paired with the corresponding biological replicate of the activity measurement. The relative phosphorylation of a residue was calculated as a ratio of number

of spectra in which the residue was identified as phosphorylated (Mascot score >32 and MD score ≥5) to number of spectra in which the residue was identified in the biological sample. The relative activity was calculated as buffer-subtracted readout of the luciferase assay relative to control. The strength of the linear association was calculated as Pearson correlation coefficient using R (packages stats) [62].

Negative stain electron microscopy

An amount of 4 µl of 0.01 mg/ml protein solution was applied to carbon-coated Cu/Pd grids and incubated for 1 min. Excess liquid was removed with the filter paper (Whatman), and the grid was washed twice with water for 5 s and 30 s. An amount of 4 µl of 2% uranyl acetate were applied twice for 5 s and 30 s. The micrographs were collected on a Talos L120C (CSSB Cryo-EM multiuser facility) with 70 000× magnification.

Bioinformatics

The homologous protein sequences were aligned with the program MAFFT [63] and depicted in Jalview. The secondary structure was predicted by Jpred [64]. The PfGSK3 structures were predicted using the Robetta server and the TrRefineRosetta method [65]. To simulate the different N-terminally phosphorylated species present in the protein fractions, we generated five models for each of the following PfGSK3 sequences: wild-type sequence, three sequences with a single phosphomimetic mutation (S40E, S42E, S43E) and one with a triple mutation S40E/S42E/S43E (abbreviated SE3). Per-residue local error estimates of the models calculated by Robetta [66] are written in the B-factor column of the PDB files. For phylogenetic analysis, protein sequences with highest identity scores were identified for individual organisms with the blastp suite of NCBI [67], the sequences were aligned with the ClustalW method and the R package msa [68], and the phylogenetic tree was calculated in PhyML with LG + G+I + F method [69]. The phylogenetic tree with multiple sequence alignment was plotted using R package ggtree [70].

Results

High yield expression and purification of enzymatically active PfGSK3 from *E. coli*

To obtain recombinant PfGSK3 in high purity and homogeneity that allow functional and structural analysis, we extended the previously established PfGSK3 expression and purification protocols [32,44,45,71]. N-terminally tagged PfGSK3 (Figure 1A) was purified by immobilized metal affinity chromatography (IMAC, Figure 1B) and subsequent size exclusion chromatography (SEC, Figure 1C). We deliberately excluded a negative IMAC step as PfGSK3 becomes sensitive to NiNTA beads upon the removal of the histidine tag and has a high tendency to precipitate. To remove impurities, we included an additional ion exchange chromatography (IEC) step (Figure 1D). This yields highly pure PfGSK3 as judged by SDS–PAGE, but the protein elutes in several partially overlapping peaks with increasing concentration of sodium chloride. Such heterogeneity of PfGSK3 might be caused by different posttranslational modifications. To assess the quality of the purified protein and to characterize it further, we applied several biophysical methods. Dynamic light scattering data confirm high homogeneity of the sample with a calculated molecular weight of 52 kDa, in agreement with the expected mass for a monomer of recombinant PfGSK3 (52.05 kDa, Supplementary Figure S2A). The secondary structure content derived from circular dichroism data is 22% α helix, 26% β sheet, 24% turns and 28% disorder. This corresponds to a typical domain organization of homologous GSK3 proteins (such as human GSK3β) with an N-terminal β-sheet-rich domain and an α-helical C-terminal domain that are flanked by disordered N-terminal and C-terminal regions (Supplementary Figure S1A, Figure 2A) [24,72]. To verify that the protein retains its three-dimensional structure and kinase properties, we tested whether its thermal stability changes in the presence of ATP and ATP analogs. Indeed, the stability of PfGSK3 markedly increased in a concentration dependent manner from 45°C to over 50°C with increasing concentration of dATP or ATP (Figure 2B and Supplementary Figure S2B). Moreover, the protein could also be stabilized by several non-hydrolysable ATP analogs (Supplementary Figure S2C). Additionally, we observed a concentration-dependent increase in the thermal stability of PfGSK3 with increasing concentrations of sulfate ions (Supplementary Figure S2D,E). Sulfate ions are bound in the activation loop of several human GSK3 structures [73–79], suggesting that the structural elements in this region are also conserved. Finally, we measured the PfGSK3 activity with a luciferase assay, using a peptide derived from human glycogen synthase 1 (GS-1) as a substrate, resulting in a catalytic rate constant of $k_{cat} = 9.5\ s^{-1}$, which is in a similar range as previously reported (between 0.7 and

PORTLAND PRESS

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

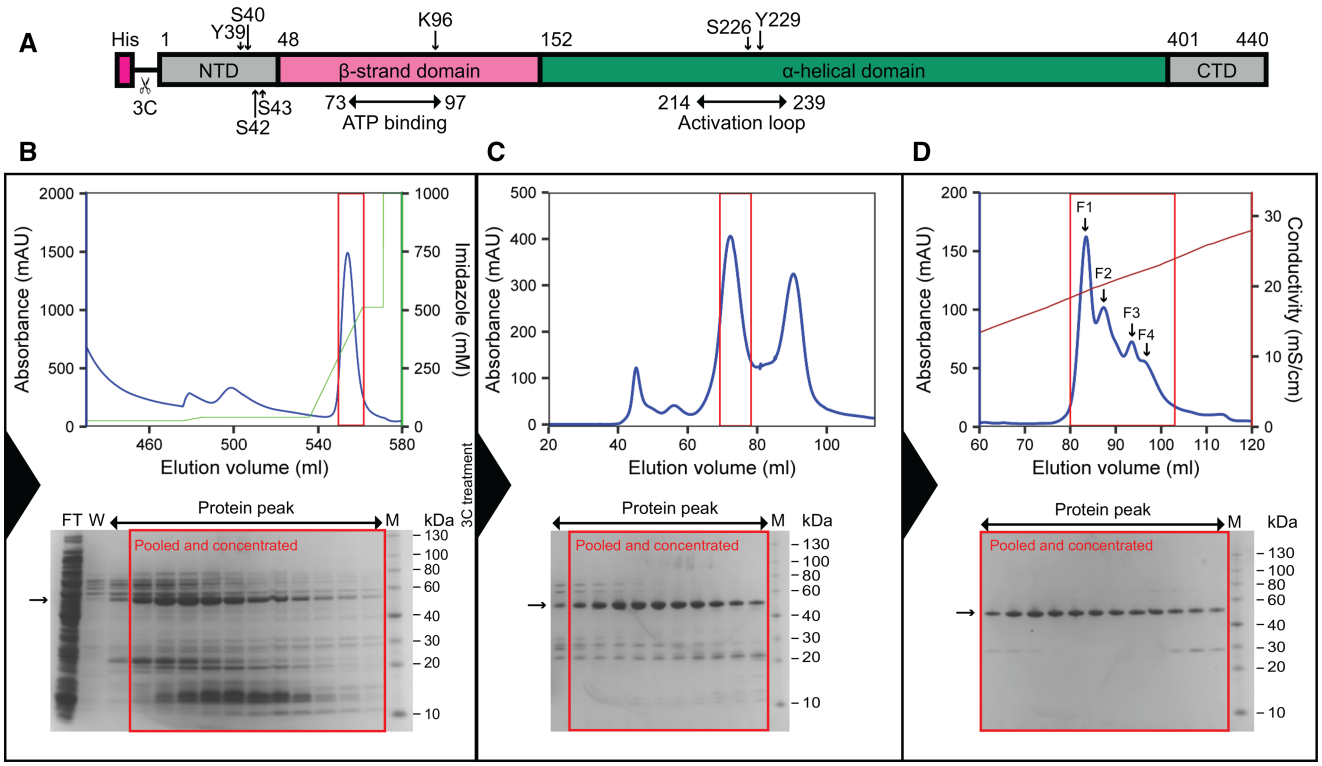


Figure 1. Expression and purification of PfGSK3.

(**A**) Construct of PfGSK3 used for expression. The construct consists of the full-length sequence of PfGSK3 with an N-terminal His-tag and a 3C cleavage site. The domain organization and phosphorylation sites are marked. NTD is the N-terminal domain, CTD is the C-terminal domain, 3C is the 3C protease cleavage cite. (**B**) First PfGSK3 purification step: imidazole gradient elution profile from the His-Trap column. (**C**) Elution profile from the Superdex 200 size exclusion column. (**D**) NaCl gradient elution profile from Resource Q ion exchange column. The peaks of the ion exchange elution represent fractions that are phosphorylated to different extent, increasing from F1 to F4. The blue curves in the chromatograms show UV absorbance, the green curve shows imidazole concentration and the brown curve shows the conductivity. The peaks in red squares in chromatograms correspond to the red squares in corresponding SDS–PAGE gels.

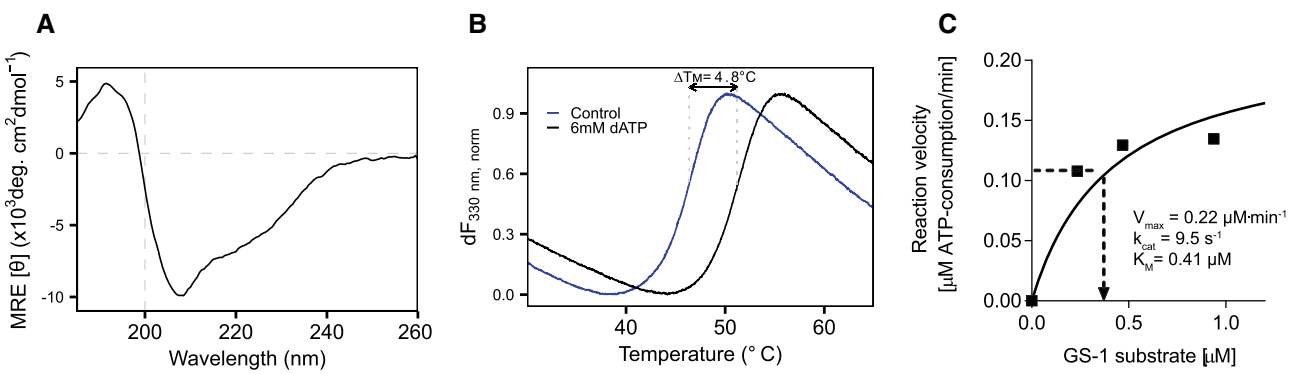


Figure 2. Characterization of PfGSK3.

(**A**) Circular dichroism shows secondary structure composition similar to what is expected for a GSK3 protein, with 22% α helix, 26% β sheet, 24% turns and 29% disordered content. The circular dichroism was measured 10× and the data were averaged, buffer subtracted, and analyzed by DichroWeb. (**B**) Thermal unfolding profiles of PfGSK3 in presence or absence of 6 mM dATP. dATP stabilizes PfGSK3, suggesting that it binds in the ATP binding pocket. (**C**) Catalytic activity of GSK3 follows Michaelis–Menten kinetics. Saturation curve shows the reaction velocity (as calculated in Supplementary Figure S2F) plotted against different substrate concentrations. Michaelis–Menten constant (K_M), maximal reaction velocity (V_{max}) and turnover number (k_{cat}) were calculated in GraphPad Prism using nonlinear regression. GS-1 = substrate peptide of the human glycogen synthase.

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

3.7 s^{-1}) for human GSK3 (Figure 2C and Supplementary Figure S2F) [72]. Thus, purified PfGSK3 is folded, stable, enzymatically active and retains the structure that is expected to be similar to human GSK3.

Divalent ions of heavy metals induce reversible formation of PfGSK3 oligomers

The observation that PfGSK3 after affinity tag cleavage has a strong tendency to precipitate upon interaction with NiNTA beads prompted us to investigate the impact of various heavy metals on PfGSK3. To this end, we measured the thermal unfolding profiles of the protein in the presence of different metal ions. Under standard conditions, PfGSK3 displays a clear transition midpoint at 45°C. However, in the presence of divalent ions, such as zinc, cobalt, nickel and copper cations, the typical unfolding transition profiles completely disappear, indicating structural changes (Supplementary Figure S3A). We selected zinc to investigate further the effect of divalent heavy metal ions on PfGSK3 because it showed the most pronounced effect on PfGSK3 unfolding. Interestingly, the secondary structure content measured by circular dichroism did not notably change upon addition of zinc, demonstrating that the secondary structural elements of the protein remain folded (Supplementary Figure S3C). However, analytical size exclusion chromatography revealed that the zinc cations induce the formation of high molecular weight (high-MW) protein species that elute earlier from the column than the metal-free protein (Figure 3A) and this effect is dependent on the concentration of the zinc cations (Supplementary Figure S3B). Indeed, in a cross-linking experiment with glutaraldehyde, the high-MW species appear predominantly in the presence of zinc (Supplementary Figure S3D). The high-MW species can revert to monomers by the addition of the divalent metal chelator EDTA (Figure 3A). To study the impact of heavy metals on the PfGSK3 function, we measured its activity in presence of zinc at different concentrations. The metal cations cause a dramatic decrease in the PfGSK3 activity (Figure 3B). However, the activity is completely restored after EDTA treatment. Finally, we used small-angle X-ray scattering (SAXS, Figure 3C and Supplementary Figure S3E) and negative stain EM (Supplementary Figure S3F) to characterize the high-MW species. The distance distribution derived from the SAXS data demonstrates high anisometry of the particles (peaking at 20 nm but with D_{max} = 90 nm). In agreement, the cross-linked high-MW species visualized by negative stain EM vary in their shape and size (Supplementary Figure S3F). In summary, heavy metal ions, such as zinc cations, induce a reversible formation of large PfGSK3 particles. While the protein maintains its secondary structure, it completely loses its enzymatic activity. Importantly, the process is reversible and the restored PfGSK3 monomers fully regain their kinase activity.

PfGSK3 exhibits autophosphorylation

PfGSK3 elutes from the IEX chromatography column as at least four different species with distinct elution volumes that are of comparable purity (Figure 1D). Such behavior could be explained by a heterogeneous

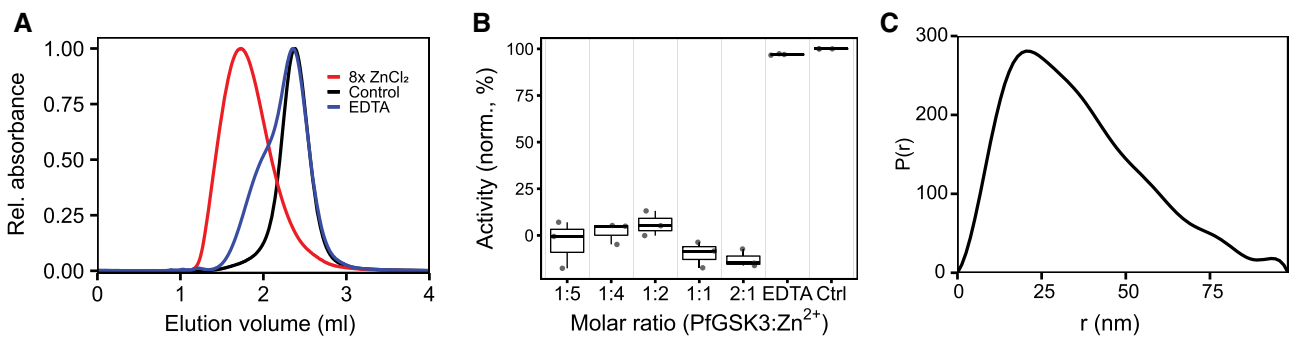


Figure 3. Heavy metal ions inhibit PfGSK3 activity.

(**A**) Analytical size exclusion chromatography profiles on a Superose 6 column of PfGSK3 apoprotein (black), after addition of 8× molar excess of zinc chloride (red) and after subsequent addition of EDTA (blue). The analysis shows that zinc ions induce the formation of high-MW PfGSK3 species. The formation is reversible because addition of EDTA shifts the elution profiles towards higher elution volumes. (**B**) The activity of PfGSK3 is strongly reduced in the presence of zinc ions, whereas the protein regains the activity after the addition of the metal chelator EDTA. (**C**) The distance distribution of the PfGSK3 high-MW species induced by zinc ions derived from SAXS data shows high heterogeneity in the sample, peaking at 20 nm but with D_{max} = 89 nm.

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

phosphorylation pattern. To investigate this possibility, we separately analyzed the four top peak fractions (F1–F4) from the IEX elution profile of the wild-type PfGSK3 preparation (Figure 1D). SDS–PAGE stained with the Pro-Q Diamond stain, which specifically binds to phosphorylated proteins, shows that the amount of total phosphorylation increases towards later elution volumes (Supplementary Figure S4A). Mass spectrometry analysis revealed that the protein is partially phosphorylated at multiple sites. We identified a number of phosphorylated residues at the N-terminus (residues 13–62, see Supplementary Figure S4B). Additionally, several residues in the activation loop, primarily S226 and Y229, but also S228 and S232 were found to be phosphorylated (Supplementary Figure S4B). The phosphorylation of the activation loop residues and two N-terminal serine residues (S31, S62) in the recombinant protein mirrors previously identified phosphorylation sites in *P. falciparum* blood-stage parasites [38–40,80–83], while the phosphorylation of other N-terminal residues, predominantly Y39, S40, S42 and S43, has not been detected in parasite lysate-based phospho-fingerprinting.

We further probed into the ability of PfGSK3 for autophosphorylation and its physiological role. First, we incubated recombinant protein with ATP/$MgCl_2$ and subsequently analysed the ion exchange chromatography profiles. Indeed, the treatment with ATP/$MgCl_2$ led to a drop in the fluorescence signal from the earlier eluting peaks and gain in the fluorescence signal from later eluting peaks, indicating that PfGSK3 is able to exhibit autophosphorylation *in vitro* (Figure 4A). Second, we employed the luciferase activity assay to examine PfGSK3 activity in the absence of a substrate peptide. The assay showed the consumption of ATP even in the absence of the substrate peptide that was dependent on the concentration of PfGSK3, confirming that PfGSK3 exhibits autophosphorylation (Figure 4B).

N-terminal domain is indispensable for PfGSK3 stability

To investigate the role of phosphorylation in PfGSK3 further, we attempted to express inactive point mutants to prevent autophosphorylation. Therefore, we changed the essential residue in the ATP binding pocket K96 or residues from the activation loop S226 and Y229 to alanine. The small-scale expression screen showed that the inactive mutants of PfGSK3 did not express in a soluble form but instead formed inclusion bodies (Figure 5A). We then isolated the inclusion bodies of the $PfGSK3^{K96A}$ and $PfGSK3^{S226A/Y229A}$ mutants and examined their phosphorylation states by mass spectrometry. These mutants display the full length sequence but are devoid of any phosphorylation (ATP binding pocket mutant K96A) or exhibit markedly reduced phosphorylation (activation loop mutant S226A/Y229A) compared with the wild-type protein (Figure 5C). As expected, these mutations drastically impair PfGSK3 function and indicate that PfGSK3 phosphorylation is important for folding

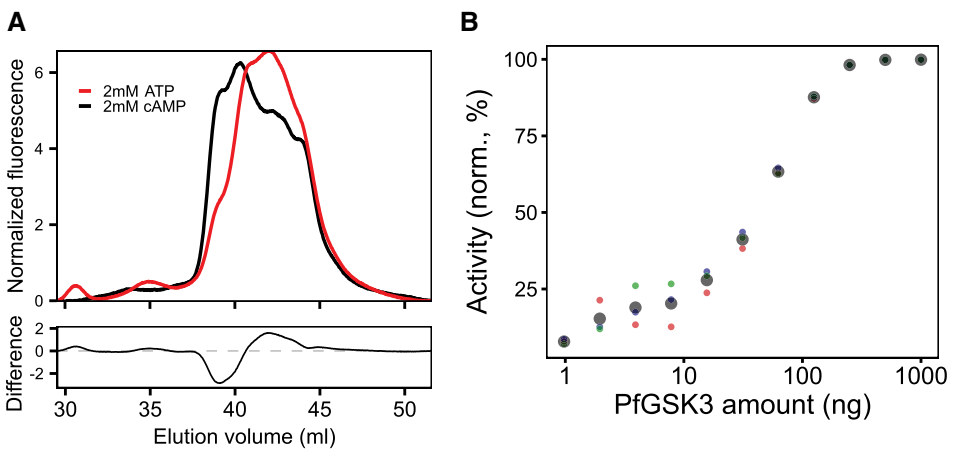


Figure 4. PfGSK3 exhibits autophosphorylation.

(**A**) IEX elution chromatograms (upper plot) of PfGSK3 after incubation with ATP (red) or cAMP as negative control (black) in the presence of magnesium ions suggest that PfGSK3 exhibits autophosphorylation. The difference between both chromatograms (bottom plot) shows an increase in later-eluting species after ATP treatment, indicating higher amount of phosphorylation. (**B**) The ATP consumption measured by the luminescence assay in the absence of the substrate is dependent on the enzyme concentration, which is indicative of the autophosphorylation exhibited by PfGSK3. The experiment was performed in triplicates. The colored points (red, green and blue) represent individual replicates, while the large black points are their average.

and solubility. The latter is further supported by a co-expression experiment of wild-type PfGSK3 with the dephosphorylating lambda phosphatase, which results in insoluble material only (Supplementary Figure S5C).

As we found the N-terminal extension of PfGSK3 to carry several phosphorylation sites that might play a role in enzyme active regulation, we cloned a series of N-terminally truncated protein constructs (Supplementary Figure S1A). None of these constructs expressed in a soluble form (Figure 5B) and mass spectrometry revealed a complete lack of phosphorylation in the PfGSK3 activation loop (ΔN64 in Figure 5C). Expression tests at different temperatures indicate that not only the aforementioned mutants and constructs, but also wild-type PfGSK3 is less expressed and mainly insoluble at lower temperatures. In fact, constructs with truncated N-termini did not express at 18°C at all, and even the wild-type protein forms only inclusion bodies under such experimental settings (Figure 5A,B). These data support the role of the PfGSK3 N-terminus and its phosphorylation for folding and function of PfGSK3.

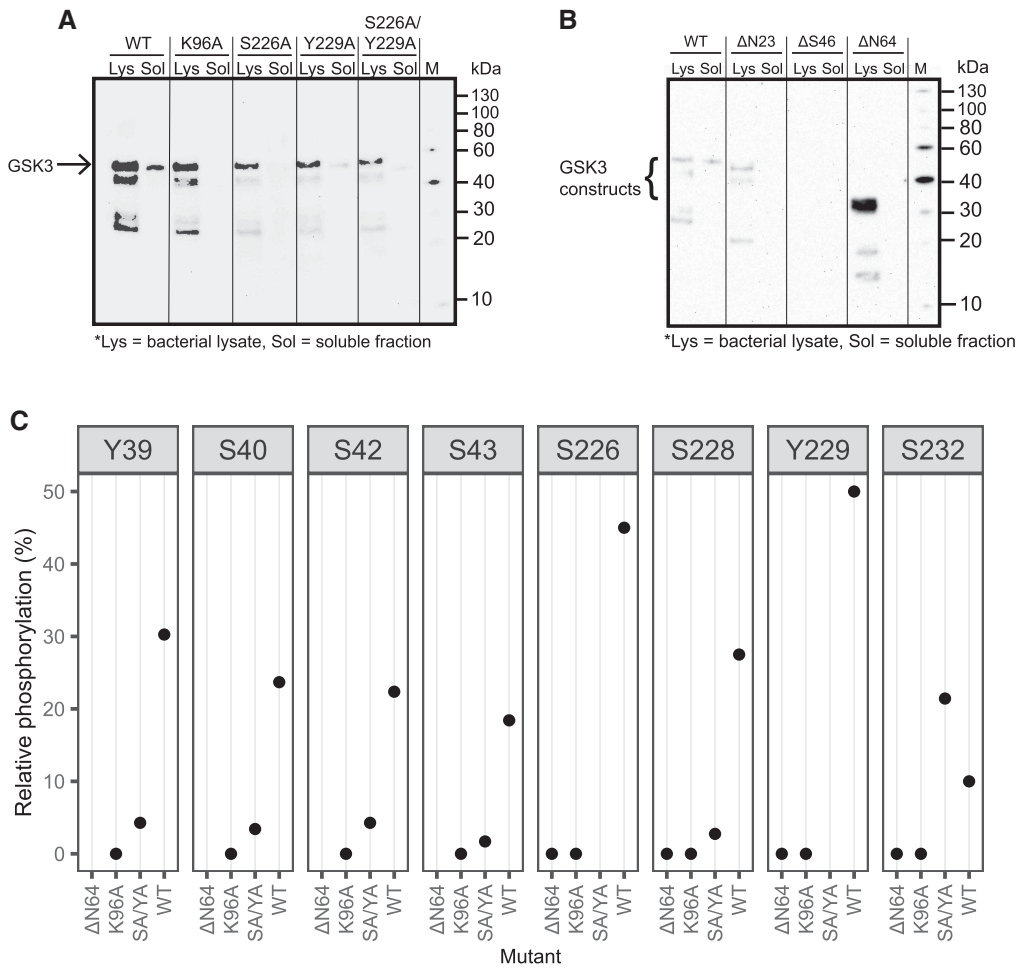


Figure 5. The N-terminus and phosphorylation are essential for PfGSK3.

(**A**) Small-scale expression test of PfGSK3 mutants with inactivating mutations in the ATP binding site (K96A) and activation loop (Y226A and S229A). Whereas all protein constructs are expressed, only the wild-type PfGSK3 is soluble, which indicates that the autophosphorylation of PfGSK3 is important for its solubility. (**B**) Small-scale expression test of PfGSK3 constructs N-terminally truncated up to the residues N23, S46 or N64, respectively. In spite of a very strong expression of PfGSK3-N64, none of the proteins is soluble, indicating that the PfGSK3 N-terminus is crucial for the autophosphorylation process. The bacterial lysates (Lys) and their soluble fractions (Sol) were analyzed by Western blot with anti-His antibodies. (**C**) Analysis of phosphorylation of PfGSK3 mutants by tandem mass spectrometry shows that the phosphorylation is completely lost (K96A, ΔN64) or reduced (S226A/Y229A, labeled here as SA/YA) compared with the wild-type protein (WT). The individual panels (from Y39 to S232) show the extent of phosphorylation in the constructs ΔN64, K96A, SA/YA and the wild-type protein (WT).

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

N-terminal phosphorylation fine-tunes the activity of PfGSK3

To further probe into the role of PfGSK3 N-terminus that might represent, like its human counterpart [24,25], the domain for enzyme activity regulation, we analyzed the individual separated fractions from IEX chromatography (Figure 1D) by LC/MS–MS. Here, we observed that the extent of phosphorylation of residues in the activation loop of PfGSK3 is comparable between the four fractions, but large differences in the N-terminal phosphorylation patterns are obvious (Figure 6A). Specifically, all N-terminal residues are significantly less phosphorylated in the first fraction (F1), while in the second fraction (F2), the extent of phosphorylation of residues Y39 and S40 is strongly increased and the third and fourth fractions (F3 and F4) have generally higher level of N-terminal phosphorylation, which is most pronounced on residues Y39, S40, S42 and S43. Of note, only residue S40 is well conserved across *Plasmodium* species, but all other *Plasmodium* species contain multiple non-conserved N-terminal residues that are potential phosphorylation sites (Supplementary Figure S6A,G).

To assess the relationship between the activity of PfGSK3 and its phosphorylation pattern, we made use of the differences in N-terminal phosphorylation between the individual IEX fractions and measured their kinase activity separately (Figure 6B). PfGSK3 activity rises with an increase in N-terminal phosphorylation levels (from F1 to F4), supporting that phosphorylation is crucial for activity. Phosphorylation levels of residues S40, S42 and S43 significantly correlate with the PfGSK3 activity (Pearson $R = 0.6$, $P < 0.05$ in all cases, Figure 6C–E), while this correlation is absent or insignificant for the other residues (Supplementary Figure S6B–F). These data support the hypothesis that phosphorylation of N-terminal PfGSK3 residues enhances PfGSK3 activity.

N-terminal phosphorylation affects the PfGSK3 structure

The phosphorylated residues at the PfGSK3 N-terminus are, based on structural homologs [24,25], expected to be distant from the active site of the protein and it is therefore unclear how N-terminal phosphorylation regulates the protein activity. In the human counterpart, the phosphorylation of S9 prompts the binding of the N-terminus to a pocket for a priming phosphorylation in the substrate-binding region and reduces the binding affinity for substrates [24,25]. Hence, we investigated structural consequences of N-terminal phosphorylation in PfGSK3. Circular dichroism data on the least (F1) and most (F4) phosphorylated fractions show no pronounced differences in their secondary structure content (Supplementary Figure S7A). Thus, to investigate the impact of the N-terminal phosphorylation on the tertiary structure of PfGSK3, we used the *de novo* structure prediction server Robetta and the TrRefineRosetta method [65] to generate initial PfGSK3 models that were further fitted against the SAXS data measured on separated PfGSK3 IEX fractions F1 and F4 (Table 1, see Figure 1D). Using the wild-type PfGSK3 sequence as well as PfGSK3 sequences with phosphomimetic mutations (see Methods), we generated 25 models (Figure 7A) that display similar structure of the core domains but are structurally diverse in both termini (Figure 7A). We identified a subset of models that provide the best fits to the experimental SAXS data (Table 1, Supplementary Figure S2B) of fractions F1 or F4 (Figure 7D). Overall, the models with a larger radius of gyration (see the point sizes in Figure 7D) display lower χ^2 (goodness-of-fit) values. Further refinement of these best-fitting extended models was conducted in the program CORAL [57] that enables assigning additional flexibility to existing models (Figure 7E,H). This approach led to poorer fits when only the C-terminus (residues 403–440) or both N-terminus and C-terminus (residues 1–64 and 403–440) were assigned flexibility (Figure 7E,H). However, conducting refinement with a disordered C-terminus (residues 403–440) and a flexible stretch of residues between the N-terminus and the PfGSK3 core (residues 47–63, conformation B, Figure 7E,H) led to two refined models displaying lower χ^2 values compared with the SAXS data from fraction F4. Ultimately, for both fractions, the models with the lowest χ^2 values display α-helical structure at the N-terminus. This α-helix is extended in the less phosphorylated F1 fraction (Figure 7B) and folds back in the more phosphorylated F4 fraction (Figure 7C). The raw SAXS data support such conformational differences, as both the calculated radius of gyration R_g and maximum intra-particle distance D_{max} decrease from F1 to F4 fraction (R_g 3.4 vs 3.2 nm; D_{max} 13 vs 12 nm, respectively, see Table 1). The dimensionless Kratky plot data representation also suggests a more extended conformation for the F1 fraction (Supplementary Figure S8C), as the F1 curve broadens and shifts to higher sR_g values relative to F4. The refined models fit the SAXS data well (χ^2 values of 1.19 and 1.06, respectively, Figure 7F,G), but should be viewed cautiously as an average representation of the conformations present in solution that contains a mixture of differently phosphorylated species. In conclusion, the combinatorial approach indeed shows that upon phosphorylation, PfGSK3 becomes more flexible, enabling a conformational change in the N-terminus. This change presumably mediates the increased PfGSK3 activity upon phosphorylation.

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

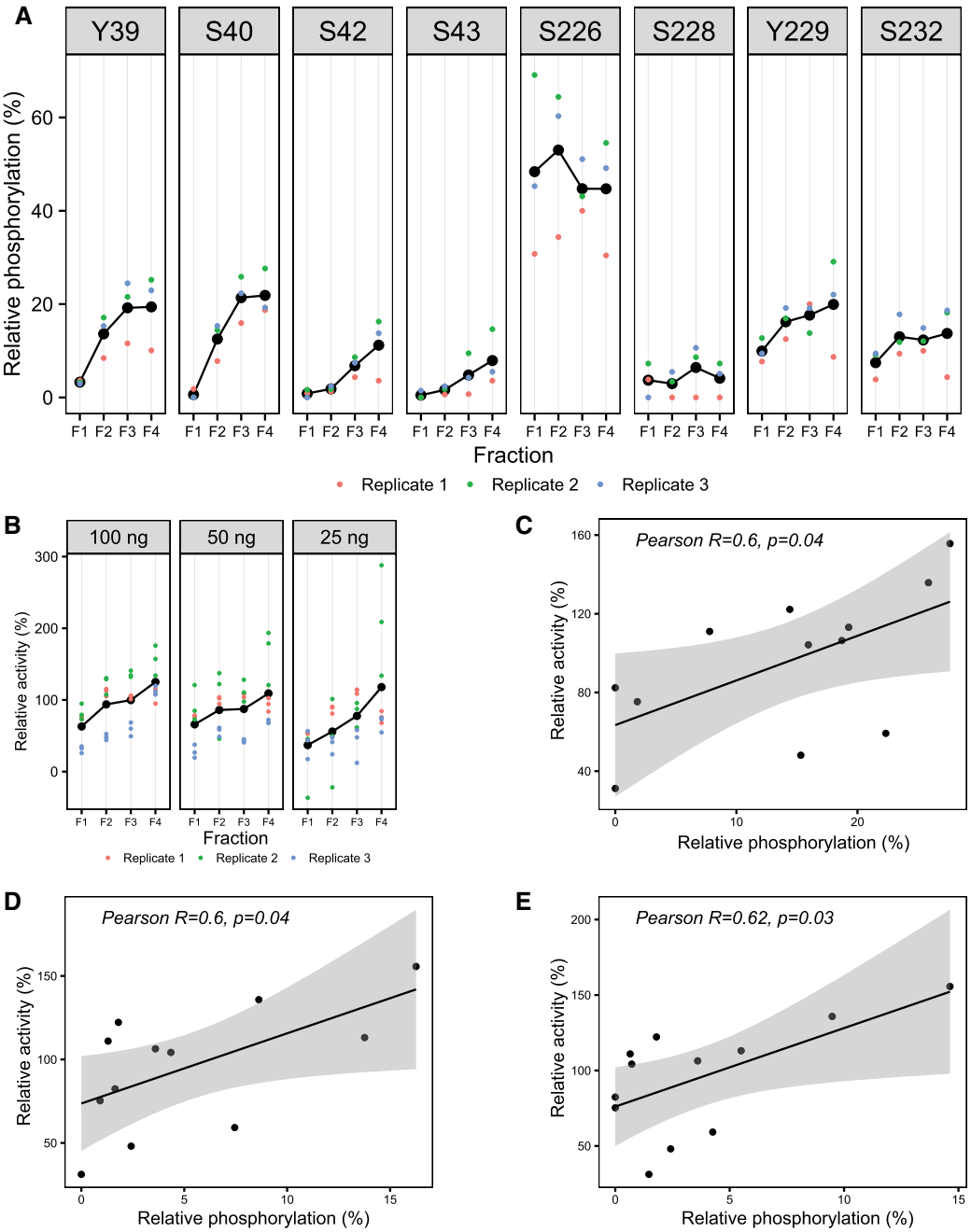


Figure 6. N-terminal phosphorylation promotes PfGSK3 activity.

(**A**) Relative phosphorylation of selected residues in the individual wild-type PfGSK3 fractions separated by ion exchange chromatography, measured by tandem mass spectrometry. The relative phosphorylation represents the fraction of all identified residues that were also identified as phosphorylated by LC–MS/MS with Mascot score > 32 and MD score ≥ 5. The data show that the residues in the N-terminal domain are gradually more phosphorylated with increasing IEX elution volume, whereas the phosphorylation at the activation loop remains constant. The different colors represent three biological triplicates. (**B**) Activity of individual PfGSK3 fractions separated by ion exchange chromatography measured with different PfGSK3 amount (100, 50 and 25 ng). The data show a general trend towards higher activity with increasing phosphorylation in the N-terminal domain. The activity was measured in biological triplicates and at different PfGSK3 concentrations. (**C**–**E**) Correlation between the relative phosphorylation of the residues S40, S42 and S43 measured by mass spectrometry, respectively, and the relative activity of the same samples (triplicates of fractions F1–F4). The data were fitted with a linear model (black line); the gray background shows the 95% confidence interval.

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

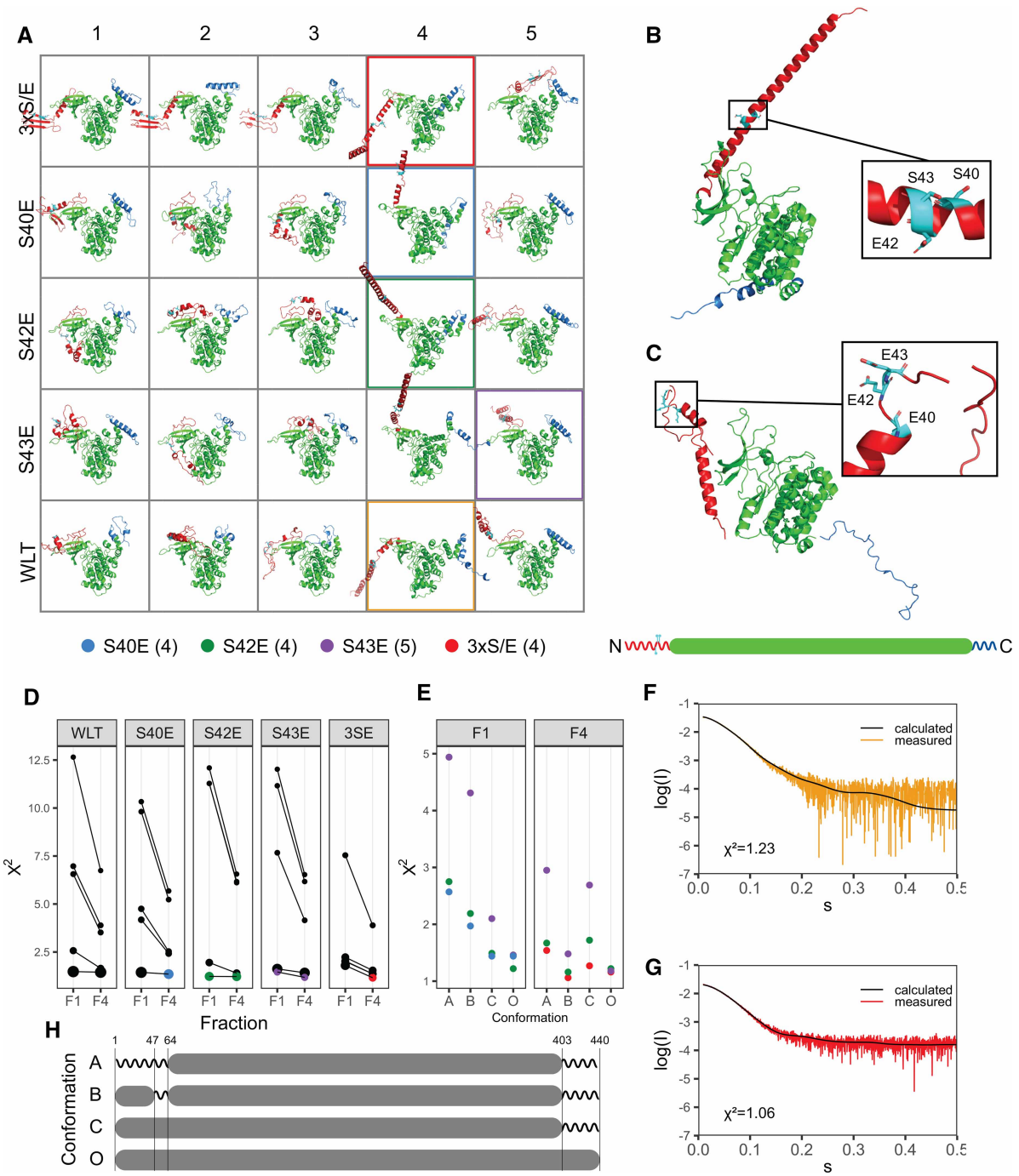


Figure 7. The N-terminus of PfGSK3 changes the structure upon phosphorylation. Part 1 of 2

(**A**) Structural models of PfGSK3 predicted by the Robetta server. Five different sequences (wild-type GSK3, mutants S40E, S42E, S43E and a triple mutant S40E/S42E/S43E, in short 3xS/E) were used as an input for the prediction and five models were predicted for each sequence. N-termini (residues 1–65) are red, residues S40, S42 and S43 cyan, the core domains green and C-termini (residues 403–440) blue. The models show a high variation in the predicted structure of N-terminus, ranging from extended to compact conformations. (**B**) The Robetta model that fits the SAXS data from F1 fraction best display an extended N-terminal helix. The inset shows the zoom of residues S40, S42 and S43 (cyan) that form a part of the N-terminal α-helix. (**C**) The Robetta model refined against SAXS data from fraction F4 using CORAL with the best fit to the data. The N-terminal helix of this model folds back towards the core of PfGSK3. The zoom in the inset shows that the phosphomimetic glutamates break the α-helix, enabling a bend that allows the N-terminal residues to fold back. (**D**) X^2 values of PfGSK3 structures modeled

Figure 7. The N-terminus of PfGSK3 changes the structure upon phosphorylation. Part 2 of 2

based on five different sequence modifications (WLT, S40E, S42E, S43E and 3xS/E) compared with SAXS data recorded for fractions F1 or F4. The identical models are connected with lines. The size of each point correlates with the radius of gyration of the particular model. Three models with lowest X^2 values for each F1 and F4 are colored (see legend). (**E**) X^2 values of the chosen models refined with CORAL using different strategies, enabling flexibility at either the N-terminus (residues 1–63, strategy A), C-terminus (residues 403–440, strategy C), or at both the C-terminus and residues 47–63 (strategy B). X^2 values of the original models are shown under 'O'. (**F**) The SAXS data measured on the PfGSK3 fraction F1 with the fit of theoretical scattering calculated from the best fitting model ($X^2 = 1.19$). (**G**) The SAXS data measured on the PfGSK3 fraction F4 with the fit of theoretical scattering calculated from the best fitting CORAL-refined model ($X^2 = 1.06$). (**H**) Graphical summary of the strategies used for model refinement against SAXS data using CORAL. The gray box symbolizes that the structure of the model was maintained and the zigzag line symbolizes that the structure was replaced with disordered residues by CORAL.

Discussion

Identification and characterization of novel drug targets in malaria remains an important goal as drug-resistant parasites are a persisting threat in the fight against malaria [2]. PfGSK3 has been identified as a potential drug target since its inhibitors also possess potent and selective anti-plasmodial activity [32,44–46]. Understanding its function is therefore important for further improvement of these inhibitors.

Here, we have shown that PfGSK3 is expressed in *E. coli* as a folded and functional protein with a heterogeneous phosphorylation pattern. We assume that the heterogeneity originating from this diverse phosphorylation hampers its detailed structural characterization by X-ray crystallography [84]. As the expression of inactive PfGSK3 protein (K96A, S226A, Y229A) as well as the removal of the heavily phosphorylated N-terminus result in insoluble protein that lacks phosphorylation, other protein constructs and strategies will likely need to be considered for detailed structural investigation of full-length PfGSK3 in the future. At the same time, these results highlight the importance of both PfGSK3 phosphorylation and the necessity of an intact N-terminus for protein stability.

Our mass spectrometry analysis of the phosphorylation pattern on recombinant PfGSK3 confirmed a number of PfGSK3 phosphorylation sites known from previous phosphoproteomics studies (S31, S62, S226, S228, Y229, S232) [38–40,80–83]. The additional phosphorylation at the N-terminus (Y39, S40, S42 and S43) described here could have occurred due to the heterologous expression, but might represent an accessory regulation site if present *in vivo*. In *h*GSK3β, the autophosphorylation of Y216 (homologous to Y229 in *P. falciparum*) in the activation loop is thought to maintain the constitutive activity of the protein [27,29], while phosphorylation of N-terminal S9 by other kinases, such as PKA and PKB, inhibits the enzyme [6,30]. Although the role of the activation loop autophosphorylation is likely conserved between the human and *P. falciparum* proteins, the role of N-terminal phosphorylation seems to be markedly different.

First, the N-terminal sequence is not conserved between the two homologous proteins, with *Plasmodium* species displaying an unusual extension with a number of potential phosphorylation sites. Second, whereas N-terminal phosphorylation of S9 in human GSK3 is mediated by other kinases, the phosphorylation of N-terminal residues in PfGSK3 is the result of autophosphorylation, as our data implies. Last, while phosphorylation of S9 in human GSK3 inhibits its activity, our work illustrates that the increasing amount of phosphorylation on the PfGSK3 N-terminus results in an increased enzymatic activity — pointing towards a different mode of kinase activity regulation.

Our structural investigation using the combination of structure prediction and its validation using experimental SAXS data provides further insights into how the N-terminal phosphorylation affects PfGSK3 structure. Two SAXS datasets originating from two differently phosphorylated IEX fractions of PfGSK3 directed the modeling towards PfGSK3 conformations with an α-helical N-terminus in both cases. However, this helix extends away from the active side of PfGSK3 in the less phosphorylated fraction, or folds back towards the center of the protein in the more phosphorylated fraction. The change in the conformation is allowed because the phosphorylation loosens the α-helical structure at the site of phosphorylation (Figure 7B) that consequently enables formation of a loop and bending of the N-terminus towards the center of the structure (Figure 7C). It is possible that the 'folded back' conformation (Figure 7C) is further stabilized by an N-terminal intramolecular coiled–coil interaction, as the probability of coiled–coil formation increases with the increasing number of N-terminal PfGSK3 phosphomimetic mutations *in silico* (Supplementary Figure S7D). Interestingly, human GSK3β undergoes a similar conformational change, where upon phosphorylation of S9, its N-terminus binds in its substrate-

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

binding region and competitively inhibits its activity [24,25]. However, N-terminal phosphorylation of PfGSK3 leads to an increase in enzymatic activity, suggesting a different regulatory mechanism. The conformational change upon phosphorylation might position the very N-terminus of PfGSK3 close to its core domains, potentially enabling intramolecular interactions that could enhance its activity.

It should be noted that the N-termini of PfGSK3 homologs are not well conserved even among the representatives of *Apicomplexa*; however, a number of residues that can be phosphorylated could still enable a similar mode of regulation as we have described here (Supplementary Figure S7A,F). Indeed, S33 of *Plasmodium berghei* GSK3 (PBANKA_0410400), equivalent to S40 in PfGSK3, was found to be phosphorylated in early gametogenesis [85]. Residues equivalent to S40 of PfGSK3 are conserved across all *Plasmodium* species and residues that can be phosphorylated are also present in the vicinity of S40 in most *Plasmodium species* (equivalent to residues 38, 39 and 42 in *P. falciparum*), suggesting that the mechanism of regulation via N-terminal phosphorylation is preserved in all *Plasmodium* GSK3 proteins.

We have shown that this N-terminal phosphorylation, as well as the phosphorylation of the activation loop, are mediated by autophosphorylation. Autophosphorylation of human GSK3β occurs directly after protein synthesis while the protein is associated with chaperones, such as HSP90 [27]. Interestingly, a major contaminant during our PfGSK3 purification process was the *E. coli* chaperone DnaK. This could indicate that this autophosphorylation also takes place directly after protein synthesis, supporting correct PfGSK3 folding.

Additionally, we have revealed another potential means of PfGSK3 regulation: divalent heavy metal ions, such as zinc, inhibit the activity of PfGSK3 by inducing the formation of heterogeneous high-MW particles. The ionic radii of these cations are in a small range of 1.09–1.21 Å, indicating that the effect is mediated through a specific binding site. The fact that the protein becomes more sensitive to the metal ions only after the cleavage of the N-terminal affinity tag indicates that this potential binding site could be located in close proximity to the N-terminus of PfGSK3. Indeed, a partially overlapping set of divalent cations inhibiting human GSK3β can be found in the literature, including lithium, beryllium, zinc, copper and mercury [17,86,87]. Lithium and beryllium ions inhibit GSK3 by competing with magnesium ions that mediate ATP binding [87–89], but to our knowledge, they have not been shown to induce the formation of high-MW particles. To this date, there has been no structure of a GSK3 homolog with a bound heavy metal ion deposited in the PDB database, and thus the inhibition mechanism of heavy metal cations remains unknown. Importantly, as the high-MW particles can dissociate back into active PfGSK3 monomers, it cannot be excluded that this type of regulation is applied *in vivo*. We have only observed incomplete dissociation of high-MW particles upon addition of excess EDTA (Figure 3A), which could indicate that the effect is partially irreversible or that the dynamics of the dissociation of high-MW particles is slow under the experimental conditions. On the other hand, metal chelating proteins and transporters were shown to affect the levels of heavy metal ions in both prokaryotes and eukaryotes [90–92] and could be utilized for regulation by *Plasmodium falciparum in vivo*. The concentration of zinc in the cytosol of infected erythrocytes peaks at the late blood stages and the intracellular levels of free zinc ions in parasites reach concentration of 1–5 mM [93], which is one order of magnitude higher than the highest concentration used in our SEC experiments (160 μM, Figure 3A) and four orders more than the highest concentration used in the kinase assay (100 nM, Figure 3B). Moreover, the concept of 'autoinhibitory polymerization' has already been described in detail for human CK2 [94], a kinase that is homologous to GSK3 [95], also exhibits autophosphorylation [96] and its *P. falciparum* counterpart displays a similar expression pattern [31], subcellular localization and function [97]. It remains to be evaluated experimentally *in vivo* whether the PfGSK3 high-MW particles play a role in its native environment.

Data Availability

The mass spectrometry data are available in the PRIDE database with dataset identifier PXD026387. The SAXS data have been deposited in the SASBDB, accession numbers SASDL77 and SASDL87. All other raw data are available on Figshare (doi: 10.6084/m9.figshare.14806887).

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

Part of this work was performed at the CryoEM Facility at CSSB, supported by the UHH and DFG grant numbers (INST 152/772-1|152/774-1|152/775-1|152/776-1|152/777-1 FUGG). This work was supported by a grant from

the Joachim Herz foundation (Nr. 800026) and the BMBF (grant number: 05K18YEA). A.A. is grateful for support by a Jürgen Manchot Stiftung fellowship.

CRediT Author Contribution

Christian Löw: Conceptualization, Resources, Supervision, Funding acquisition, Validation, Writing — original draft, Project administration, Writing — review and editing. **Samuel Pazicky:** Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. **Arne Alder:** Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. **Haydyn Mertens:** Supervision, Validation, Methodology, Writing — review and editing. **Dmitri I. Svergun:** Supervision, Validation, Investigation, Methodology, Writing — review and editing. **Tim Gilberger:** Conceptualization, Resources, Supervision, Funding acquisition, Writing — original draft, Writing — review and editing.

Acknowledgements

We thank the Sample Preparation and Characterization facility of EMBL Hamburg for support in nanoDSF, CD and DLS measurements. We also thank the Proteomics Core Facility of EMBL for the support in mass spectrometry sample preparation, measurements and analysis. We thank Maxime Killer and Vadim Kotov for the help with EM data acquisition and processing.

Abbreviations

AMA1, *Plasmodium falciparum* Apical membrane Antigen 1; GSK3β, Glycogen Synthase Kinase beta; High-MW, high molecular weight; IEX, ion exchange chromatography; IMAC, immobilized metal affinity chromatography; PfGSK3, *Plasmodium falciparum* Glycogen Synthase Kinase; SEC, size exclusion chromatography.

References

1 World Health Organization. (2019). World Malaria Report. https://www.who.int/malaria. (accessed 15 Dec. 2019). Available from: https://www.who.int/publications-detail/world-malaria-report-2019
2 Haldar, K., Bhattacharjee, S. and Safeukui, I. (2018) Drug resistance in *plasmodium*. *Nat. Rev. Microbiol.* **16**, 156–170 https://doi.org/10.1038/nrmicro.2017.161
3 Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, G. et al. (2017) A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 https://doi.org/10.1038/nrd.2016.230
4 Cabrera, D.G., Horatscheck, A., Wilson, C.R., Basarab, G., Eyermann, C.J. and Chibale, K. (2018) Plasmodial kinase inhibitors: license to cure? *J. Med. Chem.* **61**, 8061–8077 https://doi.org/10.1021/acs.jmedchem.8b00329
5 Parker, P.J., Caudwell, F.B. and Cohen, P. (1983) Glycogen synthase from rabbit skeletal muscle; effect of insulin on the state of phosphorylation of the seven phosphoserine residues *in vivo*. *Eur. J. Biochem.* **130**, 227–234 https://doi.org/10.1111/j.1432-1033.1983.tb07140.x
6 Cross, D.A.E., Alessi, D.R., Cohen, P., Andjelkovich, M. and Hemmings, B.A. (1995) Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* **378**, 785–789 https://doi.org/10.1038/378785a0
7 Welsh, G.I. and Proud, C.G. (1993) Glycogen synthase kinase-3 is rapidly inactivated in response to insulin and phosphorylates eukaryotic initiation factor eIF-2B. *Biochem. J.* **294**, 625–629 https://doi.org/10.1042/bj2940625
8 Diehl, J.A., Cheng, M., Roussel, M.F. and Sherr, C.J. (1998) Glycogen synthase kinase-3β regulates cyclin D1 proteolysis and subcellular localization. *Genes Dev.* **12**, 3499–3511 https://doi.org/10.1101/gad.12.22.3499
9 Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S. and Kikuchi, A. (1999) Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3β regulates its stability. *J. Biol. Chem.* **274**, 10681–10684 https://doi.org/10.1074/jbc.274.16.10681
10 Turenne, G.A. and Price, B.D. (2001) Glycogen synthase kinase3 beta phosphorylates serine 33 of p53 and activates p53's transcriptional activity. *BMC Cell Biol.* **2**, 12 https://doi.org/10.1186/1471-2121-2-12
11 Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S. and Kikuchi, A. (1998) Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3β and β-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. *EMBO J.* **17**, 1371–1384 https://doi.org/10.1093/emboj/17.5.1371
12 Hanger, D.P., Hughes, K., Woodgett, J.R., Brion, J.P. and Anderton, B.H. (1992) Glycogen synthase kinase-3 induces Alzheimer's disease-like phosphorylation of tau: generation of paired helical filament epitopes and neuronal localisation of the kinase. *Neurosci. Lett.* **147**, 58–62 https://doi.org/10.1016/0304-3940(92)90774-2
13 Yang, S.D., Song, J.S., Yu, J.S. and Shiah, S.G. (1993) Protein kinase FA/GSK-3 phosphorylates on Ser235-Pro and Ser404-Pro that Are abnormally phosphorylated in Alzheimer's disease brain. *J. Neurochem.* **61**, 1742–1747 https://doi.org/10.1111/j.1471-4159.1993.tb09811.x
14 Aplin, A.E., Gibb, G.M., Jacobsen, J.S., Gallo, J.M. and Anderton, B.H. (1996) *In vitro* phosphorylation of the cytoplasmic domain of the amyloid precursor protein by glycogen synthase kinase-3β. *J. Neurochem.* **67**, 699–707 https://doi.org/10.1046/j.1471-4159.1996.67020699.x
15 Ryder, J., Su, Y., Liu, F., Li, B., Zhou, Y. and Ni, B. (2003) Divergent roles of GSK3 and CDK5 in APP processing. *Biochem. Biophys. Res. Commun.* **312**, 922–929 https://doi.org/10.1016/j.bbrc.2003.11.014
16 Ferrao Santos, S., Tasiaux, B., Sindic, C. and Octave, J.N. (2011) Inhibition of neuronal calcium oscillations by cell surface APP phosphorylated on T668. *Neurobiol. Aging.* **32**, 2308–2313 https://doi.org/10.1016/j.neurobiolaging.2010.01.006

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

17 Eldar-Finkelman, H. and Martinez, A. (2011) GSK-3 Inhibitors: preclinical and clinical focus on CNS. *Front. Mol. Neurosci.* **4**, 32 https://doi.org/10.3389/fnmol.2011.00032
18 Bourel, E., Grieco, S.F. and Jope, R.S. (2015) Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. *Pharmacol. Ther.* **148**, 114–131 https://doi.org/10.1016/j.pharmthera.2014.11.016
19 Jope, R.S. (2011) Glycogen synthase kinase-3 in the etiology and treatment of mood disorders. *Front. Mol. Neurosci.* **4**, 16 https://doi.org/10.3389/fnmol.2011.00016
20 Singh, K.K. (2013) An emerging role for Wnt and GSK3 signaling pathways in schizophrenia. *Clin. Genet.* **83**, 511–517 https://doi.org/10.1111/cge.12111
21 De Sarno, P., Axtell, R.C., Raman, C., Roth, K.A., Alessi, D.R. and Jope, R.S. (2008) Lithium prevents and ameliorates experimental autoimmune encephalomyelitis. *J. Immunol.* **181**, 338–345 https://doi.org/10.4049/jimmunol.181.1.338
22 Del Ser, T., Steinwachs, K.C., Gertz, H.J., Andrés M, V., Gómez-Carrillo, B., Medina, M. et al. (2013) Treatment of Alzheimer's disease with the GSK-3 inhibitor tideglusib: a pilot study. *J. Alzheimer's Dis.* **33**, 205–215 https://doi.org/10.3233/JAD-2012-120805
23 Lovestone, S., Boada, M., Dubois, B., Hüll, M., Rinne, J.O., Huppertz, H.J. et al. (2015) A phase II trial of tideglusib in Alzheimer's disease. *J Alzheimer's Dis.* **45**, 75–88 https://doi.org/10.3233/JAD-141959
24 Haar E, T., Coll, J.T., Austen, D.A., Hsiao, H.M., Swenson, L. and Jain, J. (2001) Structure of GSK3β reveals a primed phosphorylation mechanism. *Nat. Struct. Biol.* **8**, 593–596 https://doi.org/10.1038/89624
25 Dajani, R., Fraser, E., Roe, S.M., Young, N., Good, V., Dale, T.C. et al. (2001) Crystal structure of glycogen synthase kinase 3β: structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* **105**, 721–732 https://doi.org/10.1016/S0092-8674(01)00374-9
26 Hughes, K., Nikolakaki, E., Plyte, S.E., Totty, N.F. and Woodgett, J.R. (1993) Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. *EMBO J.* **12**, 803–808 https://doi.org/10.1002/j.1460-2075.1993.tb05715.x
27 Lochhead, P.A., Kinstrie, R., Sibbet, G., Rawjee, T., Morrice, N. and Cleghon, V. (2006) A chaperone-dependent GSK3β transitional intermediate mediates activation-loop autophosphorylation. *Mol. Cell* **24**, 627–633 https://doi.org/10.1016/j.molcel.2006.10.009
28 Cole, A.R., Causeret, F., Yadirgi, G., Hastie, C.J., McLauchlan, H., McManus, E.J. et al. (2006) Distinct priming kinases contribute to differential regulation of collapsin response mediator proteins by glycogen synthase kinase-3 *in vivo*. *J. Biol. Chem.* **281**, 16591–16598 https://doi.org/10.1074/jbc.M513344200
29 Cole, A.R., Knebel, A., Morrice, N.A., Robertson, L.A., Irving, A.J., Connolly, C.N. et al. (2004) GSK-3 phosphorylation of the Alzheimer epitope within collapsin response mediator proteins regulates axon elongation in primary neurons. *J. Biol. Chem.* **279**, 50176–50180 https://doi.org/10.1074/jbc.C400412200
30 Sutherland, C., Leighton, I.A. and Cohen, P. (1993) Inactivation of glycogen synthase kinase-3β by phosphorylation: new kinase connections in insulin and growth-factor signalling. *Biochem. J.* **2296**, 15–19 https://doi.org/10.1042/bj2960015
31 Wichers, J.S., Scholz, J.A.M., Strauss, J., Witt, S., Lill, A., Ehnold, L.I. et al. (2019) Dissecting the gene expression, localization, membrane topology, and function of the plasmodium falciparum STEVOR protein family. *MBio* **10**, e01500-19 https://doi.org/10.1128/mBio.01500-19
32 Prinz, B., Harvey, K.L., Wilcke, L., Ruch, U., Engelberg, K., Biller, L. et al. (2016) Hierarchical phosphorylation of apical membrane antigen 1 is required for efficient red blood cell invasion by malaria parasites. *Sci Rep.* **6**, 34479 https://doi.org/10.1038/srep34479
33 Mordue, D.G., Desai, N., Dustin, M. and Sibley, L.D. (1999) Invasion by *Toxoplasma gondii* establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. *J. Exp. Med.* **190**, 1783–1792 https://doi.org/10.1084/jem.190.12.1783
34 Cowman, A.F., Tonkin, C.J., Tham, W.H. and Duraisingh, M.T. (2017) The molecular basis of erythrocyte invasion by malaria parasites. *Cell Host Microbe* **22**, 232–245 https://doi.org/10.1016/j.chom.2017.07.003
35 Narum, D.L. and Thomas, A.W. (1994) Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Mol. Biochem. Parasitol.* **67**, 59–68 https://doi.org/10.1016/0166-6851(94)90096-5
36 Triglia, T., Healer, J., Caruana, S.R., Hodder, A.N., Anders, R.F., Crabb, B.S. et al. (2000) Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species. *Mol. Microbiol.* **38**, 706–718 https://doi.org/10.1046/j.1365-2958.2000.02175.x
37 Treeck, M., Zacherl, S., Herrmann, S., Cabrera, A., Kono, M., Struck, N.S. et al. (2009) Functional analysis of the leading malaria vaccine candidate AMA-1 reveals an essential role for the cytoplasmic domain in the invasion process. *PLoS Pathog.* **5**, e1000322 https://doi.org/10.1371/journal.ppat.1000322
38 Treeck, M., Sanders, J.L., Elias, J.E. and Boothroyd, J.C. (2011) The phosphoproteomes of plasmodium falciparum and toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe* **10**, 410–419 https://doi.org/10.1016/j.chom.2011.09.004
39 Solyakov, L., Halbert, J., Alam, M.M., Semblat, J.-P., Dorin-Semblat, D., Reininger, L. et al. (2011) Global kinomic and phospho-proteomic analyses of the human malaria parasite plasmodium falciparum. *Nat. Commun.* **2**, 565 https://doi.org/10.1038/ncomms1558
40 Lasonder, E., Green, J.L., Camarda, G., Talabani, H., Holder, A.A., Langsley, G. et al. (2012) The plasmodium falciparum schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase A signaling. *J. Proteome Res.* **11**, 5323–5337 https://doi.org/10.1021/pr300557m
41 Leykauf, K., Treeck, M., Gilson, P.R., Nebl, T., Braulke, T., Cowman, A.F. et al. (2010) Protein kinase a dependent phosphorylation of apical membrane antigen 1 plays an important role in erythrocyte invasion by the malaria parasite. *PLoS Pathog.* **6**, e1000941 https://doi.org/10.1371/journal.ppat.1000941
42 Patel, A., Perrin, A.J., Flynn, H.R., Bisson, C., Withers-Martinez, C., Treeck, M. et al. (2019) Cyclic AMP signalling controls key components of malaria parasite host cell invasion machinery. *PLoS Biol.* **17**, e3000264 https://doi.org/10.1371/journal.pbio.3000264
43 Wilde, M.L., Triglia, T., Marapana, D., Thompson, J.K., Kouzmitchev, A.A., Bullen, H.E. et al. (2019) Protein kinase A is essential for invasion of *Plasmodium falciparum* into human erythrocytes. *MBio* **10**, e01972-19 https://doi.org/10.1128/mBio.01972-19
44 Fugel, W., Oberholzer, A.E., Gschloessl, B., Dzikowski, R., Pressburger, N., Preu, L. et al. (2013) 3,6-Diamino-4-(2-halophenyl)-2-benzoylthieno[2,3-b]pyridine-5-carbonitriles are selective inhibitors of plasmodium falciparum glycogen synthase kinase-3. *J. Med. Chem.* **56**, 264–275 https://doi.org/10.1021/jm301575n
45 Masch, A., Nasereddin, A., Alder, A., Bird, M.J., Schweda, S.I., Preu, L. et al. (2019) Structure-activity relationships in a series of antiplasmodial thieno[2,3-b]pyridines. *Malar. J.* **18**, 1–10 https://doi.org/10.1186/s12936-019-2725-y
46 Moolman, C., van der Sluis, R., Beteck, R.M. and Legoabe, L.J. (2021) Exploration of benzofuran-based compounds as potent and selective *Plasmodium falciparum* glycogen synthase kinase-3 (PfGSK-3) inhibitors. *Bioorg. Chem.* **112**, 104839 https://doi.org/10.1016/j.bioorg.2021.104839

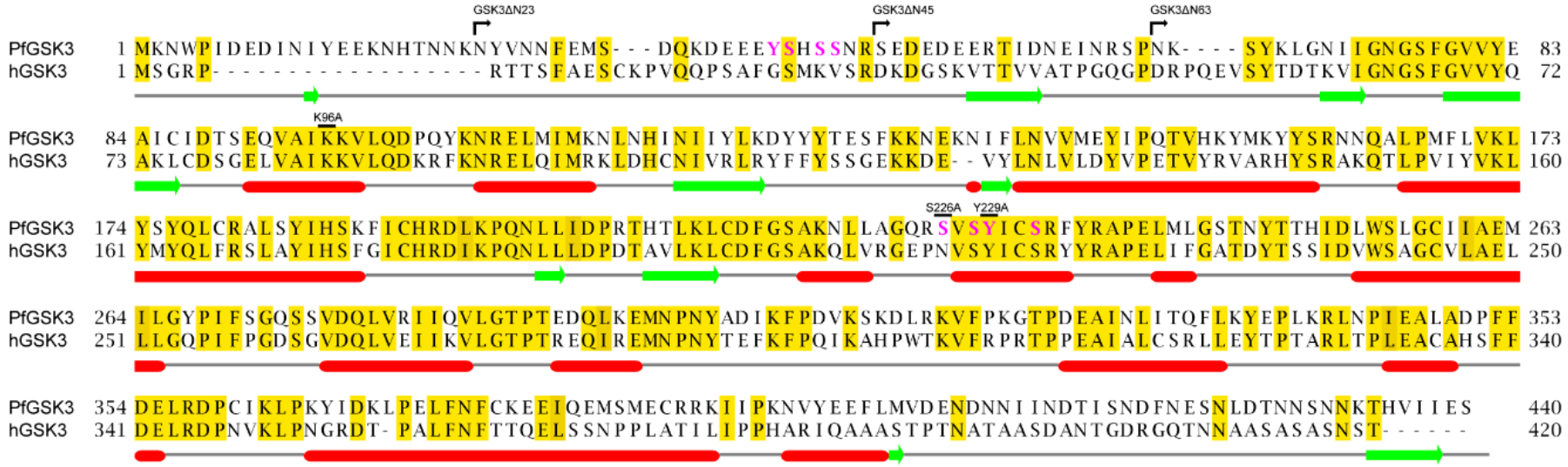
47 Osolodkin, D.I., Zakharevich, N.V., Palyulin, V.A., Danilenko, V.N. and Zefirov, N.S. (2011) Bioinformatic analysis of glycogen synthase kinase 3: human versus parasite kinases. *Parasitology* **138**, 725–735 https://doi.org/10.1017/S0031182011000151
48 Berrow, N.S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N. et al. (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35**, e45 https://doi.org/10.1093/nar/gkm047
49 Zhang, Y., Werling, U. and Edelmann, W. (2014) Seamless ligation cloning extract (SLiCE) cloning method. *Methods Mol. Biol.* **1116**, 235–244 https://doi.org/10.1007/978-1-62703-764-8_16
50 Manavalan, P. and Johnson, W.C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76–85 https://doi.org/10.1016/0003-2697(87)90135-7
51 Sreerama, N. and Woody, R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260 https://doi.org/10.1006/abio.2000.4880
52 Whitmore, L. and Wallace, B.A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392–400 https://doi.org/10.1002/bip.20853
53 Blanchet, C.E., Zozulya A, V., Kikhney, A.G., Franke, D., Konarev P, V., Shang, W. et al. (2012) Instrumental setup for high-throughput small-and wide-angle solution scattering at the X33 beamline of EMBL hamburg. *J. Appl. Crystallogr.* **45**, 489–495 https://doi.org/10.1107/S0021889812013490
54 Graewert, M.A., Franke, D., Jeffries, C.M., Blanchet, C.E., Ruskule, D., Kuhle, K. et al. (2015) Automated pipeline for purification, biophysical and X-ray analysis of biomacromolecular solutions. *Sci. Rep.* **5**, 10734 https://doi.org/10.1038/srep10734
55 Franke, D., Petoukhov M, V., Konarev P, V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T. et al. (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 https://doi.org/10.1107/S1600576717007786
56 Svergun, D., Barberato, C. and Koch, M.H. (1995) CRYSOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 https://doi.org/10.1107/S0021889895007047
57 Petoukhov M, V., Franke, D., Shkumatov A, V., Tria, G., Kikhney, A.G., Gajda, M. et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 https://doi.org/10.1107/S0021889812007662
58 Schweda, S.I., Alder, A. and Gilberger, T.) 4-arylthieno[2,3-b]pyridine-2-carboxamides are a new class of antiplasmodial agents sandra. *Molecules* **25**, 3187 https://doi.org/10.3390/molecules25143187
59 Sridharan, S., Kurzawa, N., Werner, T., Günthner, I., Helm, D., Huber, W. et al. (2019) Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat. Commun.* **10**, 1155 https://doi.org/10.1038/s41467-019-09107-y
60 Franken, H., Mathieson, T., Childs, D., Sweetman, G.M.A., Werner, T., Tögel, I. et al. (2015) Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* **10**, 1567–1593 https://doi.org/10.1038/nprot.2015.101
61 Savitski, M.M., Wilhelm, M., Hahne, H., Kuster, B. and Bantscheff, M. (2015) A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteomics* **14**, 2394–2404 https://doi.org/10.1074/mcp.M114.046995
62 R Core Team (2020) *R: A Language and Environment for Statistical Computing. R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria
63 Rozewicki, J., Li, S., Amada, K.M., Standley, D.M. and Katoh, K. (2019) MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* **47**, W5–W10 https://doi.org/10.1093/nar/gky874
64 Cole, C., Barber, J.D. and Barton, G.J. (2008) The jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–W201 https://doi.org/10.1093/nar/gkn238
65 Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S. and Baker, D. (2020) Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl Acad. Sci. U.S.A.* **117**, 1496–1503 https://doi.org/10.1073/pnas.1914677117
66 Hiranuma, N., Park, H., Baek, M., Anishchenko, I., Dauparas, J. and Baker, D. (2021) Improved protein structure refinement guided by deep learning based accuracy estimation. *Nat Commun.* **12**, 1340 https://doi.org/10.1038/s41467-021-21511-x
67 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 https://doi.org/10.1186/1471-2105-10-421
68 Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C. and Hochreiter, S. (2015) Msa: an R package for multiple sequence alignment. *Bioinformatics* **31**, 3997–3999 https://doi.org/10.1093/bioinformatics/btv494
69 Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* **59**, 307–321 https://doi.org/10.1093/sysbio/syq010
70 Yu, G. (2020) Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinformatics* **35**, 3041–3043 https://doi.org/10.1002/cpbi.96
71 Droucheau, E., Primot, A., Thomas, V., Mattei, D., Knockaert, M., Richardson, C. et al. (2004) Plasmodium falciparum glycogen synthase kinase-3: molecular model, expression, intracellular localisation and selective inhibitors. *Biochim. Biophys. Acta Proteins Proteomics* **1697**, 181–196 https://doi.org/10.1016/j.bbapap.2003.11.023
72 Dajani, R., Fraser, E., Roe, S.M., Yeo, M., Good, V.M., Thompson, V. et al. (2003) Structural basis for recruitment of glycogen synthase kinase 3β to the axin-APC scaffold complex. *EMBO J.* **22**, 494–501 https://doi.org/10.1093/emboj/cdg068
73 Henley, Z.A., Bax, B.D., Inglesby, L.M., Champigny, A., Gaines, S., Faulder, P. et al. (2017) From PIM1 to PI3Kδ via GSK3β: target hopping through the kinome. *ACS Med. Chem. Lett.* **8**, 1093–1098 https://doi.org/10.1021/acsmedchemlett.7b00296
74 Gentile, G., Merlo, G., Pozzan, A., Bernasconi, G., Bax, B., Bamborough, P. et al. (2012) 5-Aryl-4-carboxamide-1,3-oxazoles: potent and selective GSK-3 inhibitors. *Bioorg. Med. Chem. Lett.* **22**, 1989–1994 https://doi.org/10.1016/j.bmcl.2012.01.034
75 Gentile, G., Bernasconi, G., Pozzan, A., Merlo, G., Marzorati, P., Bamborough, P. et al. (2011) Identification of 2-(4-pyridyl)thienopyridinones as GSK-3β inhibitors. *Bioorg. Med. Chem. Lett.* **21**, 4823–4827 https://doi.org/10.1016/j.bmcl.2011.06.050
76 Chu, N., Salguero, A.L., Liu, A.Z., Chen, Z., Dempsey, D.R., Ficarro, S.B. et al. (2018) Akt kinase activation mechanisms revealed using protein semisynthesis. *Cell* **174**, 897–907.e14 https://doi.org/10.1016/j.cell.2018.07.003
77 Bax, B., Carter, P.S., Lewis, C., Guy, A.R., Bridges, A., Tanner, R. et al. (2001) The structure of phosphorylated GSK-3β complexed with a peptide, FRATtide, that inhibits β-catenin phosphorylation. *Structure* **9**, 1143–1152 https://doi.org/10.1016/S0969-2126(01)00679-7

Biochemical Journal (2022) **479** 337–356
https://doi.org/10.1042/BCJ20210829

78 Liang, S.H., Chen, J.M., Normandin, M.D., Chang, J.S., Chang, G.C., Taylor, C.K. et al. (2016) Discovery of a highly selective glycogen synthase kinase-3 inhibitor (PF-04802367) that modulates tau phosphorylation in the brain: translation for PET neuroimaging. *Angew. Chem. Int. Ed. Engl.* **55**, 9601–9605 https://doi.org/10.1002/anie.201603797
79 Wang, Y., Wach, J.Y., Sheehan, P., Zhong, C., Zhan, C., Harris, R. et al. (2016) Diversity-oriented synthesis as a strategy for fragment evolution against GSK3β. *ACS Med. Chem. Lett.* **7**, 852–856 https://doi.org/10.1021/acsmedchemlett.6b00230
80 Pease, B.N., Huttlin, E.L., Jedrychowski, M.P., Talevich, E., Harmon, J., Dillman, T. et al. (2013) Global analysis of protein expression and phosphorylation of three stages of plasmodium falciparum intraerythrocytic development. *J. Proteome Res.* **12**, 4028–4045 https://doi.org/10.1021/pr400394g
81 Ganter, M., Goldberg, J.M., Dvorin, J.D., Paulo, J.A., King, J.G., Tripathi, A.K. et al. (2017) *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony. *Nat. Microbiol.* **2**, 17017 https://doi.org/10.1038/nmicrobiol.2017.17
82 Kumar, S., Kumar, M., Ekka, R., Dvorin, J.D., Paul, A.S., Madugundu, A.K. et al. (2017) PfCDPK1 mediated signaling in erythrocytic stages of *Plasmodium falciparum*. *Nat. Commun.* **8**, 63 https://doi.org/10.1038/s41467-017-00053-1
83 Flueck, C., Drought, L.G., Jones, A., Patel, A., Perrin, A.J., Walker, E.M. et al. (2019) Phosphodiesterase beta is the master regulator of camp signalling during malaria parasite invasion. *PLoS Biol.* **17**, e3000154 https://doi.org/10.1371/journal.pbio.3000154
84 Shrestha, A., Hamilton, G., O'Neill, E., Knapp, S. and Elkins, J.M. (2012) Analysis of conditions affecting auto-phosphorylation of human kinases during expression in bacteria. *Protein Expr. Purif.* **81**, 136–143 https://doi.org/10.1016/j.pep.2011.09.012
85 Invergo, B.M., Brochet, M., Yu, L., Choudhary, J., Beltrao, P. and Billker, O. (2017) Sub-minute phosphoregulation of cell cycle systems during plasmodium gamete formation. *Cell Rep.* **21**, 2017–2029 https://doi.org/10.1016/j.celrep.2017.10.071
86 Ilouz, R., Kaidanovich, O., Gurwitz, D. and Eldar-Finkelman, H. (2002) Inhibition of glycogen synthase kinase-3β by bivalent zinc ions: insight into the insulin-mimetic action of zinc. *Biochem. Biophys. Res. Commun.* **295**, 102–106 https://doi.org/10.1016/S0006-291X(02)00636-8
87 Ryves, W.J., Harwood, A.J., Dajani, R. and Pearl, L. (2002) Glycogen synthase kinase-3 inhibition by lithium and beryllium suggests the presence of two magnesium binding sites. *Biochem. Biophys. Res. Commun.* **290**, 967–972 https://doi.org/10.1006/bbrc.2001.6305
88 Ryves, W.J. and Harwood, A.J. (2001) Lithium inhibits glycogen synthase kinase-3 by competition for magnesium. *Biochem. Biophys. Res. Commun.* **280**, 720–725 https://doi.org/10.1006/bbrc.2000.4169
89 Mudireddy, S.R., Abdul, A.R.M., Gorjala, P. and Gary, R.K. (2014) Beryllium is an inhibitor of cellular GSK-3β that is 1,000-fold more potent than lithium. *BioMetals* **27**, 1203–1216 https://doi.org/10.1007/s10534-014-9783-y
90 Choveaux, D.L., Przyborski, J.M. and Goldring, J.P.D. (2012) A *Plasmodium falciparum* copper-binding membrane protein with copper transport motifs. *Malar. J.* **11**, 397 https://doi.org/10.1186/1475-2875-11-397
91 Cheng, T., Xia, W., Wang, P., Huang, F., Wang, J. and Sun, H. (2013) Histidine-rich proteins in prokaryotes: metal homeostasis and environmental habitat-related occurrence. *Metallomics* **5**, 1423–1429 https://doi.org/10.1039/c3mt00059a
92 Navarro, J.A. and Schneuwly, S. (2017) Copper and zinc homeostasis: lessons from *Drosophila melanogaster*. *Front. Genet.* **8**, 223 https://doi.org/10.3389/fgene.2017.00223
93 Marvin, R.G., Wolford, J.L., Kidd, M.J., Murphy, S., Ward, J., Que, E.L. et al. (2012) Fluxes in 'free' and total zinc are essential for progression of intraerythrocytic stages of *Plasmodium falciparum*. *Chem. Biol.* **19**, 731–741 https://doi.org/10.1016/j.chembiol.2012.04.013
94 Lolli, G., Pinna, L.A. and Battistutta, R. (2012) Structural determinants of protein kinase CK2 regulation by autoinhibitory polymerization. *ACS Chem. Biol.* **7**, 1158–1163 https://doi.org/10.1021/cb300054n
95 Tewari, R., Straschil, U., Bateman, A., Böhme, U., Cherevach, I., Gong, P. et al. (2010) The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission. *Cell Host Microbe* **8**, 377–387 https://doi.org/10.1016/j.chom.2010.09.006
96 Ruiz-Carrillo, D., Lin, J., El Sahili, A., Wei, M., Sze, S.K., Cheung, P.C.F. et al. (2018) The protein kinase CK2 catalytic domain from *Plasmodium falciparum*: crystal structure, tyrosine kinase activity and inhibition. *Sci. Rep.* **8**, 7365 https://doi.org/10.1038/s41598-018-25738-5
97 Engelberg, K., Paul, A.S., Prinz, B., Kono, M., Ching, W., Heincke, D. et al. (2013) Specific phosphorylation of the PfRh2b invasion ligand of *Plasmodium falciparum*. *Biochem. J.* **452**, 457–466 https://doi.org/10.1042/BJ20121694

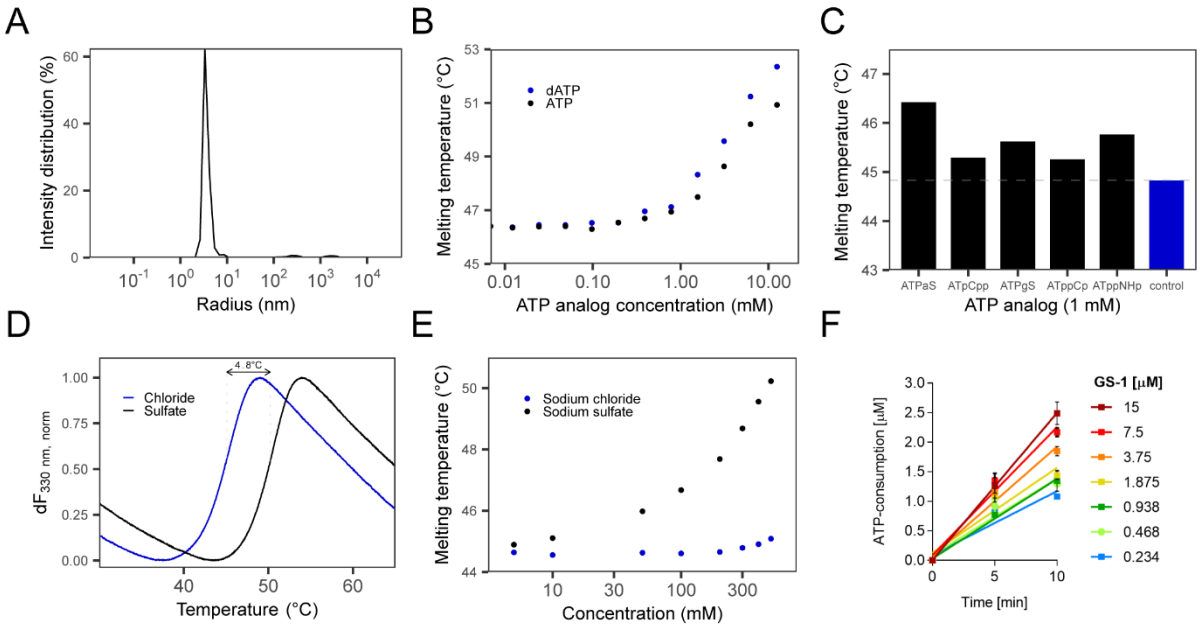
Supporting information

Supplementary figure 1



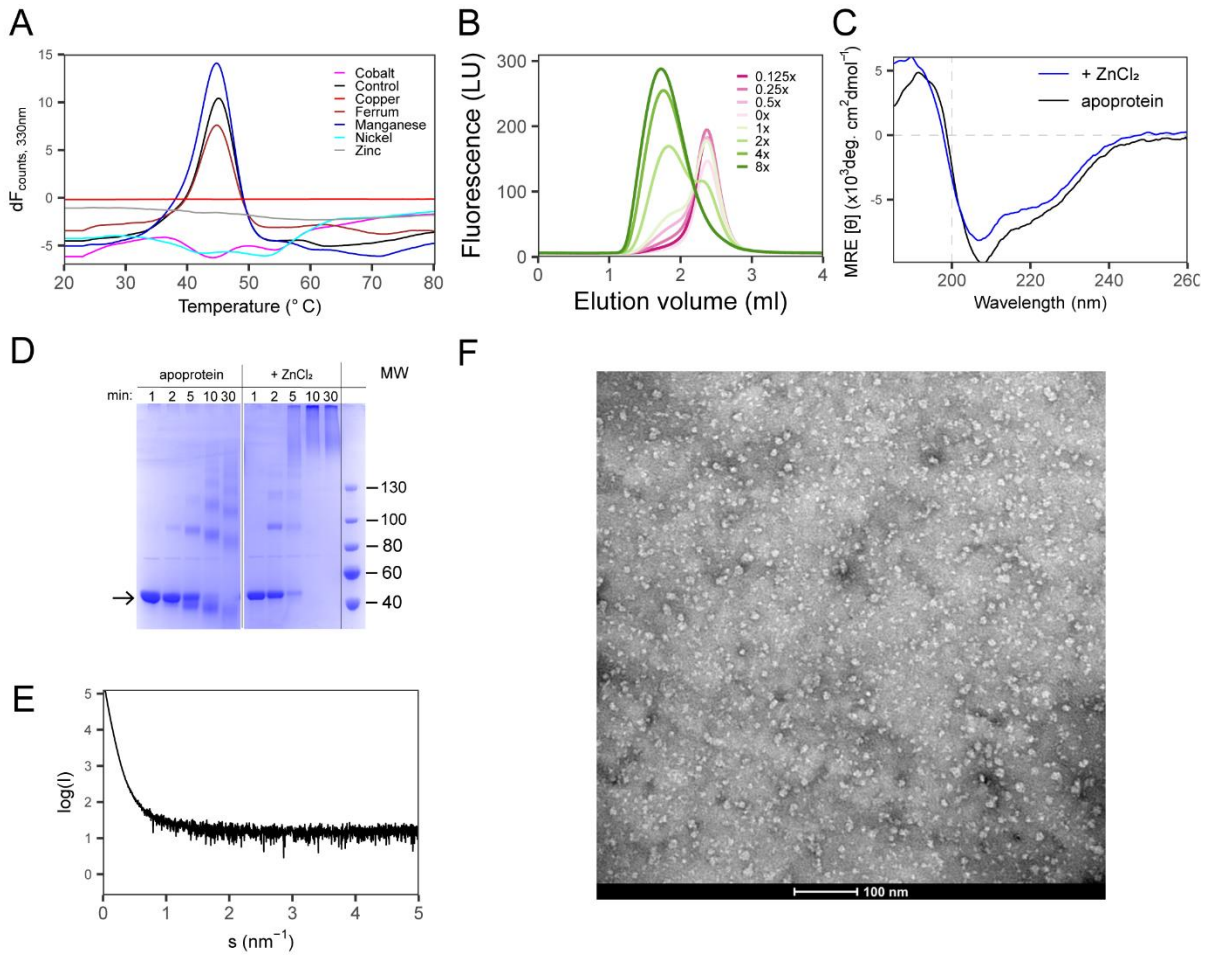
Sequence alignment of PfGSK3 with human GSK3. The conserved residues are highlighted in yellow. Residues mutated in this study and the N-truncated constructs are indicated. Residues that are phosphorylated are shown in magenta. The PfGSK3 secondary structure prediction by JPred (98) is indicated below the sequence. The sequences were aligned using MAFFT (63) and plotted in Jalview (64).

Supplementary figure 2



(A) Analysis of dynamic light scattering data of PfGSK3 reveals a high degree of homogeneity with a molecular weight estimated from the measured radius equal to the expected value of 52 kDa. (B) Stability dependence of the PfGSK3 melting temperature measured by nano differential scanning fluorimetry (nDSF) in the presence of different concentrations of ATP or dATP. Melting temperature increases with the concentration of ATP or dATP, indicative for binding ATP and its analogs. (C) Melting temperature of PfGSK3 measured by nDSF with or without 1 mM non-hydrolyzable ATP analogs. (D) Thermal unfolding profiles of PfGSK3 in presence of 500 mM sodium chloride or sodium sulfate show that PfGSK3 is preferentially stabilized by sulfate ions. (E) Stabilization of PfGSK3 by sulfate ions is concentration-dependent, indicating that stabilization effect is caused by specific sulfate ion binding. (F) Initial reaction velocity of GSK3 kinase reaction with 20 ng kinase, 6 µM ATP and substrate concentrations as indicated (GS-1 = peptide substrate of the human glycogen synthase). ATP consumption was quantified using the KinaseGlo assay (Promega) 5 and 10 min after start of kinase reaction. Lines were generated in GraphPad Prism using linear regression. Slope of each line represents reaction velocity (ATP consumption/min). Data points represent mean ± SD of three kinase reactions.

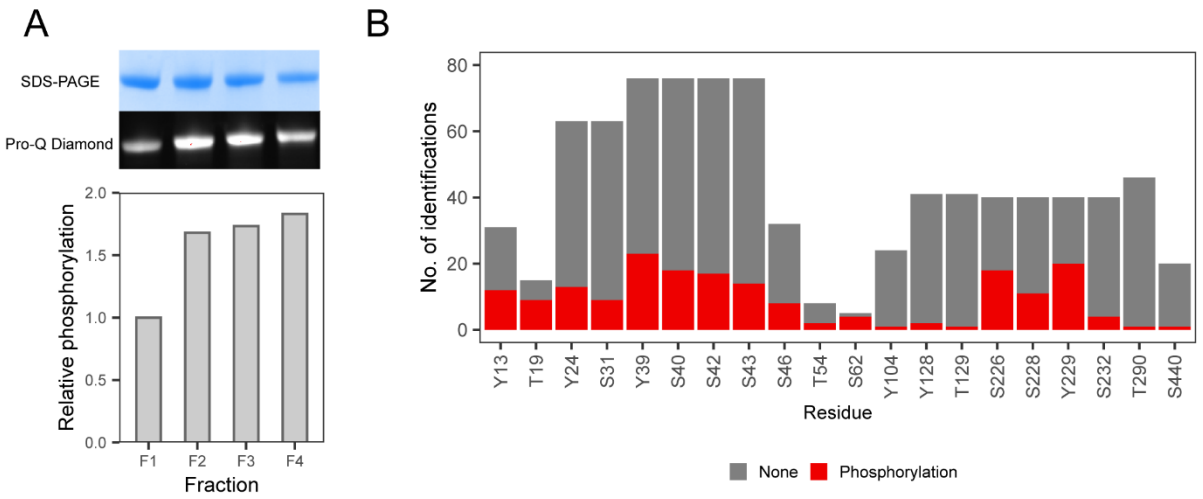
Supplementary figure 3



(A) Effect of heavy metals on thermal unfolding of PfGSK3. The unfolding trace was measured by nano differential scanning fluorimetry at a protein concentration of 1 mg/ml and a metal concentration of 1 mM. Presence of some heavy metal ions alters the unfolding profile of PfGSK3. (B) Effect of zinc chloride on the elution profile of PfGSK3 on a Superose 6 column. The experiment shows that the zinc ions induce the formation of high-MW particles in a concentration-dependent manner. The numbers indicate the molar access of metal over the protein. (C) Circular dichroism spectrum of PfGSK3 in the presence of zinc shows that the secondary structure is retained even in the induced high-MW particles. (D) Cross-linking experiment confirm that compared to apoprotein, PfGSK3 forms high-MW species that are induced by addition of zinc ions. (E) SAXS curve of PfGSK3 in the presence of zinc shows that the high-MW particles are large and heterogeneous. (F) Negative stain electron

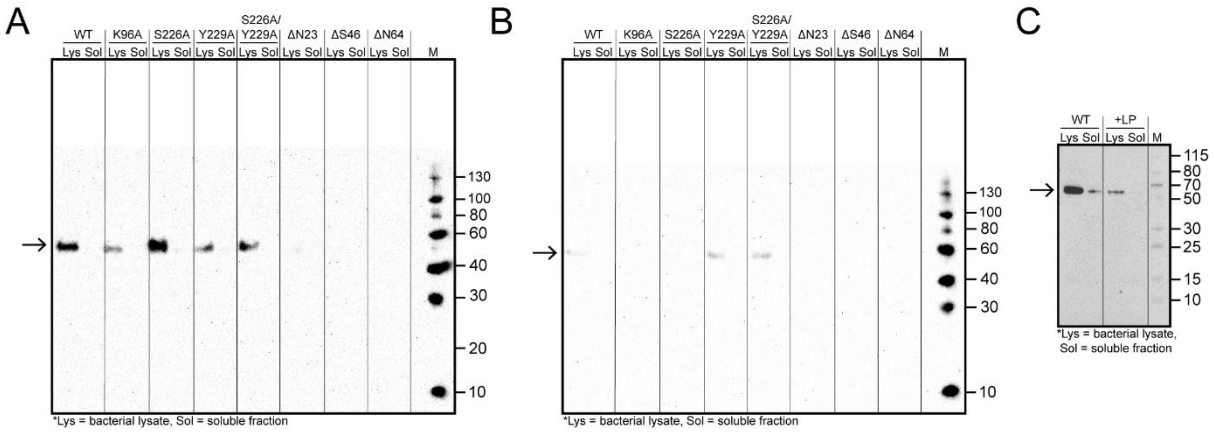
microscopy image of PfGSK3 high-MW particles induced by copper further stress the large sample heterogeneity.

Supplementary figure 4



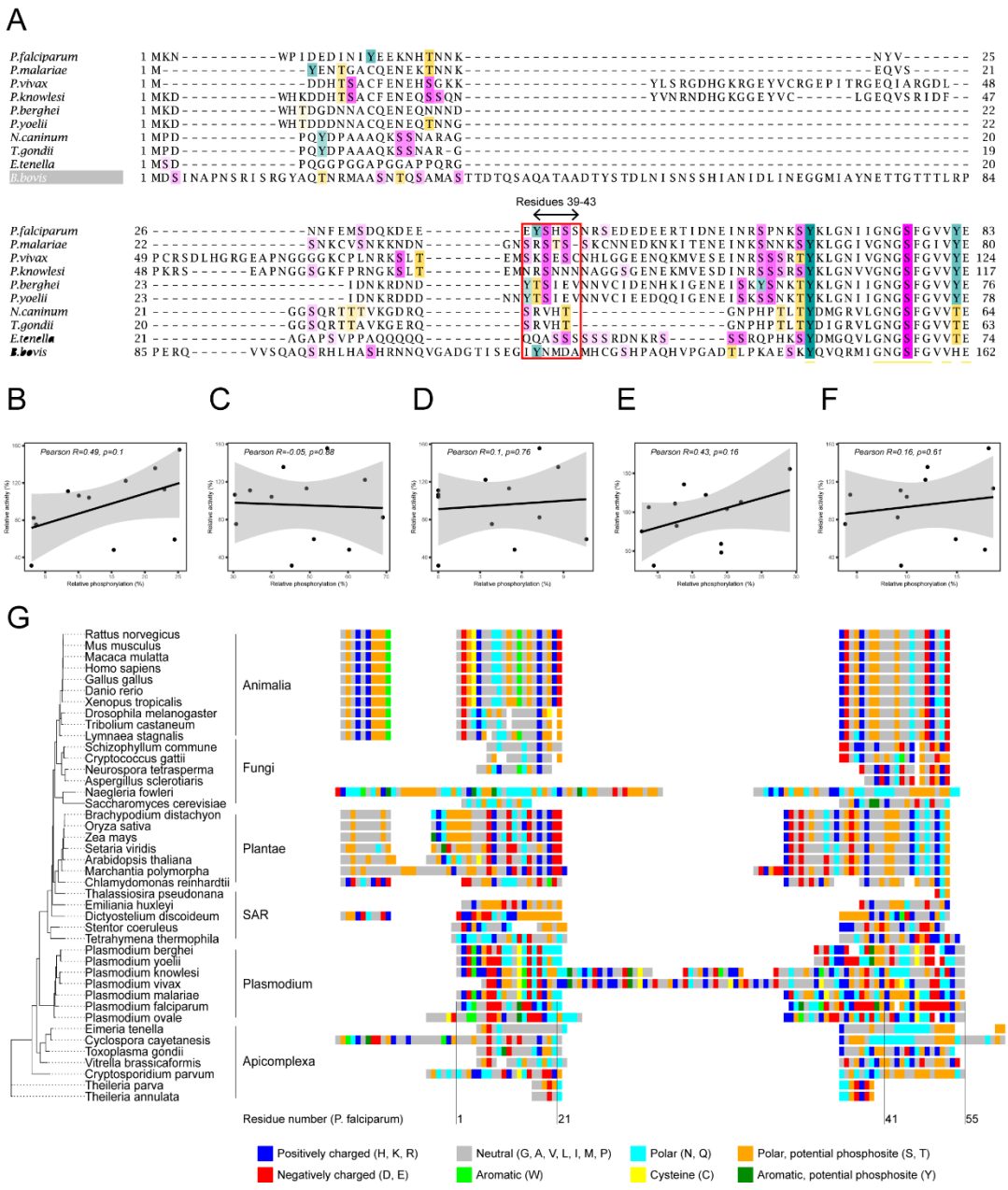
(A) The intensity ratio of bands between SDS-PAGE and Pro-Q diamond staining for individual IEX fraction shows that the later eluting species of PfGSK3 are more phosphorylated. The relative phosphorylation was calculated as ratio of relative band intensities with F1 band intensities as reference values. (B) LC MS-MS analysis of the PfGSK3 phosphorylation sites reveals heterogeneous phosphorylation of PfGSK3. All residues identified to be phosphorylated are shown.

Supplementary figure 5



(A-B) Western blot analysis of expression test of PfGSK3 mutants in its ATP binding site (K96) or activation loop (S226 and Y229) and of PfGSK3 N-terminally truncated constructs at 25°C (A) and 18°C (B). The results show that the lower temperatures decrease the expression level and solubility of the proteins. (C) Western blot of test of PfGSK3 co-expression with λ-phosphatase shows that PfGSK3 dephosphorylation renders the protein insoluble.

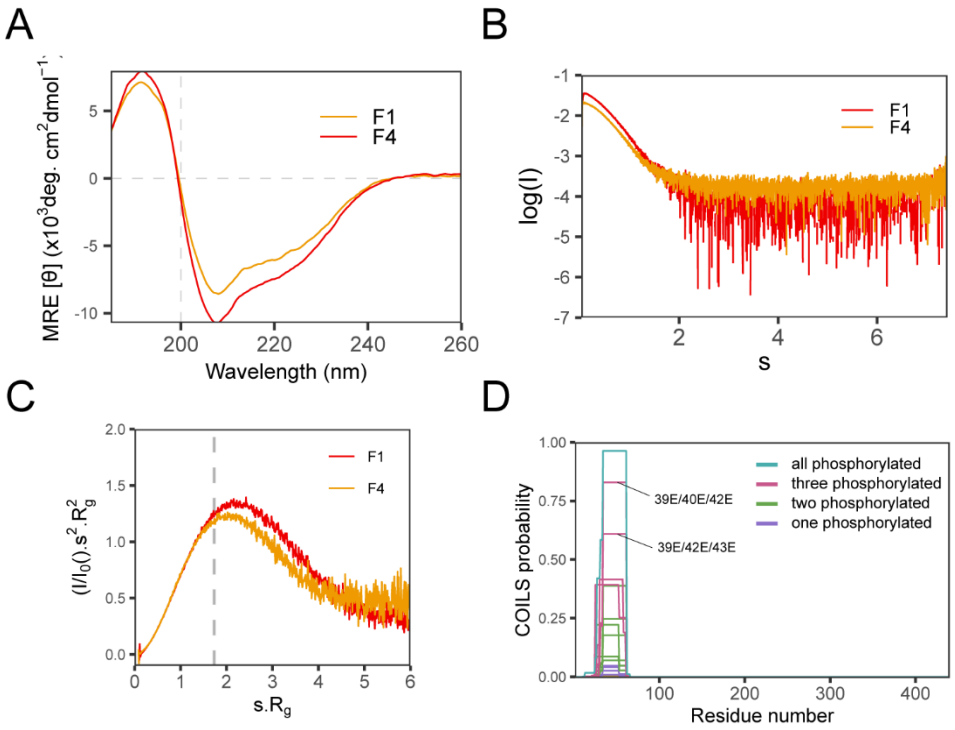
Supplementary figure 6



(A) Sequence alignment of PfGSK3 with homologous proteins from other *Plasmodia* and *Apicomplexa*. Only N-terminal residues are shown. Serine, threonine and tyrosine residues are depicted with magenta, yellow and cyan background, respectively. The region around regulatory N-terminal residues is highlighted in a red box. The alignment shows that residue S40 and neighboring residues are conserved in *Plasmodium* but not within *Apicomplexa*. (B-F) Correlation between the protein activity and the phosphorylation of residues Y39, S226, S228, Y229 and Y232, respectively. (G) Phylogenetic analysis of GSK3 shows that the N-terminus

is unique in *Plasmodium* species as they separate into a clade distinct from other apicomplexans. Abbreviation SAR refers to a super group that includes Stramenopiles, Alveolates and Rhizaria.

Supplementary figure 7



(A) Circular dichroism data measured on PfGSK3 fractions F1 and F4 indicates that the secondary structure does not change upon phosphorylation. (B) Small angle X-ray scattering data measured on PfGSK3 fractions F1 and F4. (C) Dimensionless Kratky plot representation of the SAXS data measured on fractions F1 and F4 shows that the phosphorylation of PfGSK3 (F4) leads to a less extended structural conformation. (D) *In silico* prediction of coiled-coil formation using the COILS server with PfGSK3 sequences with N-terminal residues (Y39, S40, S42 and S43) mutated to phosphomimetic glutamate residues used as an input.