1   **N-terminal phosphorylation regulates the activity of Glycogen Synthase Kinase 3**

2   **from *Plasmodium falciparum***

3

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

5   Christian Löw[1,2*]

6

7   [1] Centre for Structural Systems Biology (CSSB), Notkestrasse 85, D-22607 Hamburg,

8   Germany.

9   [2] Molecular Biology Laboratory (EMBL), Hamburg Unit c/o Deutsches Elektronen

10   Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany.

11   [3] Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Strasse 74, D-20359

12   Hamburg, Germany.

13   [4] Department of Biology, University of Hamburg, Hamburg, Germany

14

15   *Corresponding author

16   Christian Löw

17   European Molecular Biology Laboratory Hamburg, Notkestrasse 85, D-22607 Hamburg,

18   Germany.

19   Fax: +49 4089902149

20   Tel: +49 408998-87570

21   e-mail: christian.loew@embl-hamburg.de

22   https://orcid.org/0000-0003-0764-7483

23   **Running title:** N-terminal phosphorylation in PfGSK3

24   **Keywords:** malaria, GSK3, autophosphorylation, drug target, SAXS

25

26 **Abstract**

27 As the decline of malaria cases stalled over the last five years, novel targets in *Plasmodium*

28 *falciparum* are necessary for the development of new drugs. Glycogen Synthase Kinase

29 (PfGSK3) has been identified as a potential target, since its selective inhibitors were shown to

30 disrupt the parasite`s life cycle. Here, we show that PfGSK3 exhibits autophosphorylation,

31 leading to an extensive phosphorylation both *in vitro* and in the parasite. In the uncanonical

32 N-terminal region of the parasite enzyme, we identified several autophosphorylation sites that

33 regulate the activity of PfGSK3. By combining molecular modeling with experimental small-

34 angle X-ray scattering data, we show that increased PfGSK3 activity is promoted by

35 conformational changes in the PfGSK3 N-terminus, triggered by N-terminal phosphorylation.

36 Our work provides novel insights into the structure and regulation of the malarial PfGSK3.

2

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 are encoding 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). The invasion process is an essential, parasite-driven event that relies on the orchestrated interplay of multiple protein-protein interactions (33). It requires the apical membrane antigen 1 (AMA1), a type I transmembrane protein located in the micronemes of the parasite and translocated on its surface before the host cell entry (32, 34–39). The N-terminal ectodomain of AMA1 interacts with the rhoptry neck (RON) protein complex of the

3

64   parasite, that is inserted into the membrane of the red blood cell, forming the basis of the so

65   called "tight junction" between the membrane of the invading parasite and its target cell (40–

66   44). The short C-terminus of AMA1 locates inside the parasite and thus, is amenable to

67   regulation by phosphorylation (45, 46). In a two-step phosphorylation event, protein kinase A

68   (PfPKA) first phosphorylates S610 of AMA1 (45, 47, 48), which in turn enables subsequent

69   phosphorylation of AMA1 residue T613 by PfGSK3 (32). Thus, PfGSK3 serves as a secondary

70   kinase, although both phosphorylation sites are equally important for AMA1 function and *P.*

71   *falciparum* invasion (32). PfGSK3 inhibitors, built on a thieno[2,3-b]pyridine or benzofuran

72   scaffold, have been developed that display selective antiplasmodial activity with IC_{50} values in

73   the low micromolar to sub-micromolar range (49–51) and one of these inhibitors has been

74   shown to abrogate the red blood cell invasion (32). Further improvements of these inhibitors

75   could be guided by a three-dimensional structure of PfGSK3 that has, however, not been

76   determined so far.

77   Based on sequence homology across GSK3 proteins, PfGSK3 consists of two conserved

78   structural domains: a β-sheet-rich lobe with an ATP binding pocket and an α-helical lobe with

79   the activation loop and a substrate binding site. These domains are typically preceded by an

80   N-terminal domain that is predicted to be disordered (25). In PfGSK3, the important residues

81   in the ATP binding pocket and in the activation loop remain conserved, however, its

82   N-terminus displays an unique extension and its function remains unknown (52).

83   To analyze this in more detail, we optimized the purification protocol of PfGSK3 that enabled

84   us to perform thorough structural and functional characterization of the protein *in vitro*. We

85   show evidence that PfGSK3 exhibits autophosphorylation of the N-terminal and activation loop

86   residues *in vitro* and validated these phosphorylation sites *in vivo*. Phosphorylation of the

87   N-terminal residues induces conformational changes of the protein that lead to an increase in

88   the enzymatic activity of PfGSK3, providing a new regulatory mechanism.

4

Results

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

89

90

91 To obtain recombinant PfGSK3 in high purity and homogeneity that allow functional and

92 structural analysis, we extended the previously established PfGSK3 expression and

93 purification protocols (32, 49, 50, 53). In short, N-terminally hexahistidine-tagged PfGSK3 (Fig.

94 1A) was separated from the bacterial cell lysate by immobilized metal affinity chromatography

95 (IMAC, Fig. 1B) and subsequent size exclusion chromatography (SEC, Fig. 1C). We

96 deliberately excluded a negative IMAC step as PfGSK3 becomes sensitive to NiNTA beads

97 upon the removal of the histidine tag and has a high tendency to precipitate. Instead, to remove

98 the impurities, we included an additional ion exchange chromatography (IEC) step (Fig. 1D).

99 This yields the desired level of purity of PfGSK3 as judged by SDS-PAGE, but the protein

100 elutes in several partially overlapping peaks with increasing concentration of sodium chloride.

101 Such heterogeneity of PfGSK3 might be caused by different posttranslational modifications.

102 To assess the quality of the purified protein and to characterize it further, we applied several

103 biophysical methods. Dynamic light scattering data confirm high homogeneity of the sample

104 with a calculated molecular weight of 52 kDa, in agreement with the expected mass for a

105 monomer of recombinant PfGSK3 (52.05 kDa, Fig. S2A). The secondary structure content

106 derived from circular dichroism data is 22% α helix, 26% β sheet, 24% turns and 28% disorder.

107 This corresponds to a typical domain organization of homologous GSK3 proteins (such as

108 human GSK3β) with an N-terminal β-sheet-rich domain and an α-helical C-terminal domain

109 that are flanked by disordered N-terminal and C-terminal regions (Fig. S1A, Fig. 2A) (24, 54).

110 To verify that the protein retains its three-dimensional structure and kinase properties, we

111 tested whether its thermal stability changes in the presence of ATP and ATP analogs. Indeed,

112 the stability of PfGSK3 markedly increased in a concentration dependent manner from 45°C

113 to over 50°C with increasing concentration of dATP or ATP (Fig. 2B and Fig. S2B). Moreover,

114 the protein could also be stabilized by several non-hydrolysable ATP analogs (Fig. S2C).

115 Additionally, we observed a concentration-dependent increase of the thermal stability of

116   PfGSK3 with increasing concentrations of sulfate ions (Fig. S2D-E). Sulfate ions are bound in

117   the activation loop of several human GSK3 structures (55–61), suggesting that the structural

118   elements in this region are also conserved. Finally, we measured the PfGSK3 activity with a

119   luciferase assay, using a peptide derived from human glycogen synthase 1 (GS-1) as a

120   substrate, resulting in a catalytic rate constant of k_{cat} = 9.5 s^{-1}, which is in a similar range as

121   previously reported (between 0.7 and 3.7 s^{-1}) for human GSK3 (Fig. 2C, Fig. S2F) (54). Thus,

122   purified PfGSK3 is folded, stable, enzymatically active and retains the structure that is

123   expected to be similar to human GSK3.

124   **Divalent ions of heavy metals induce reversible formation of PfGSK3 oligomers**

125   The observation that PfGSK3 after affinity tag cleavage has a strong tendency to precipitate

126   upon interaction with NiNTA beads prompted us to investigate the impact of various heavy

127   metals on PfGSK3. To this end, we measured the thermal unfolding profiles of the protein in

128   the presence of different metal ions. Under standard conditions, PfGSK3 displays a clear

129   transition midpoint at 45°C. However, in the presence of divalent ions, such as zinc, cobalt,

130   nickel and copper cations, the typical unfolding transition profiles completely disappear,

131   indicating structural changes (Fig. S3A). We selected zinc to investigate further the effect of

132   divalent heavy metal ions on PfGSK3 because it showed the most pronounced effect on

133   PfGSK3 unfolding. Interestingly, the secondary structure content measured by circular

134   dichroism did not notably change upon addition of zinc, demonstrating that the secondary

135   structural elements of the protein remain folded (Fig. S3C). However, analytical size exclusion

136   chromatography revealed that the zinc cations induce the formation of high molecular weight

137   (high-MW) protein species that elute earlier from the column than the metal-free protein (Fig.

138   3A) and this effect is dependent on the concentration of the zinc cations (Fig. S3B). Indeed,

139   in a cross-linking experiment with glutaraldehyde, the high-MW species appear predominantly

140   in the presence of zinc (Fig. S3D). The high-MW species can revert to monomers by the

141   addition of the divalent metal chelator EDTA (Fig. 3A). To study the impact of heavy metals

142   on the PfGSK3 function, we measured its activity in presence of zinc at different

143   concentrations. The metal cations cause a dramatic decrease in the PfGSK3 activity (Fig. 3B).

144   However, the activity is completely restored after EDTA treatment. Finally, we used

145   small-angle X-ray scattering (SAXS, Fig. 3C and Fig. S3E) and negative stain EM (Fig. S3F)

146   to characterize the high-MW species. The distance distribution derived from the SAXS data

147   demonstrates high anisometry of the particles (peaking at 20 nm but with D_{max}=90 nm). In

148   agreement, the cross-linked high-MW species visualized by negative stain EM vary in their

149   shape and size (Fig. S3F). In summary, heavy metal ions, such as zinc cations, induce a

150   reversible formation of large PfGSK3 particles. While the protein maintains its secondary

151   structure, it completely loses its enzymatic activity. Importantly, the process is reversible and

152   the restored PfGSK3 monomers fully regain their kinase activity.

153   **PfGSK3 exhibits autophosphorylation**

154   PfGSK3 elutes from the IEX chromatography column in at least four different species with

155   distinct elution volumes that are of comparable purity (Fig. 1D). Such behavior could be

156   explained by a heterogeneous phosphorylation pattern. To investigate this possibility, we

157   separately analyzed the four top peak fractions (F1-F4) from the IEX elution profile of the wild

158   type PfGSK3 preparation (Fig. 1D). SDS-PAGE stained with the Pro-Q Diamond stain, which

159   specifically binds to phosphorylated proteins, shows that the amount of total phosphorylation

160   increases towards later elution volumes (Fig. S4A). Mass spectrometry analysis revealed that

161   the protein is partially phosphorylated at multiple sites. We identified a number of

162   phosphorylated residues at the N-terminus (residues 13-62, see Fig. S4B). Additionally,

163   several residues in the activation loop, primarily S226 and Y229, but also S228 and S232 were

164   found to be phosphorylated (Fig. S4B). The phosphorylation of the activation loop residues

165   and two N-terminal serine residues (S31, S62) is in agreement with the previously identified

166   phosphorylation sites in *P. falciparum* blood-stage parasites (37–39, 62–65), while the

167   phosphorylation of other N-terminal residues, predominantly Y39, S40, S42 and S43, has not

168   been detected in parasite lysate-based phospho-fingerprinting. As the *E. coli* expression

169   system only possesses a limited capacity to phosphorylate proteins (66–68), we tested the

7

170   ability of PfGSK3 to exhibit autophosphorylation by incubating recombinant protein with

171   ATP/MgCl$_2$ and subsequent analysis of the ion exchange chromatography profiles. Indeed,

172   the treatment with ATP/MgCl$_2$ led to a drop in the fluorescence signal from the earlier eluting

173   peaks and gain in the fluorescence signal from later eluting peaks, indicating that PfGSK3 is

174   able to exhibit autophosphorylation *in vitro* (Fig. 4A). In addition, we employed the luciferase

175   activity assay to examine PfGSK3 activity in the absence of a substrate peptide. The assay

176   showed the consumption of ATP even in the absence of the substrate peptide that was

177   dependent on the concentration of PfGSK3, confirming that PfGSK3 exhibits

178   autophosphorylation (Fig. 4B).

The N-terminus of endogenous PfGSK3 is phosphorylated

180   Having observed autophosphorylation of PfGSK3 *in vitro*, we also investigated the

181   phosphorylation state of PfGSK3 expressed in the parasite. To this end, we made use of a

182   *P. falciparum* cell line endogenously expressing C-terminally GFP-tagged GSK3 (GSK3-GFP,

183   Fig. S5). The parasites were harvested at the schizont stage and GSK3-GFP was purified

184   using GFP Trap agarose beads (Fig. 5A) and subjected to mass spectrometry analysis. The

185   identified peptides exhibited phosphorylation both at the N-terminus and in the activation loop,

186   similarly to recombinantly expressed PfGSK3 (Fig. 5B). The generally low Mascot Delta Score

187   (69) did not allow us to confidently pinpoint the exact phosphorylation sites, but confirmed that

188   a significant portion of the residues on the peptide derived from the unstructured N-terminus

189   were phosphorylated (average 28%, residues 39-43) compared to the activation loop peptides

190   (average 5%, residues 226-233). These data attest that N-terminal phosphorylation of PfGSK3

191   also occurs *in vivo.*

N-terminal domain is indispensable for PfGSK3 stability

193   To investigate the role of phosphorylation in PfGSK3 further, we attempted to express inactive

194   mutants to prevent autophosphorylation. Therefore, we changed the essential residue in the

195   ATP binding pocket K96 or residues from the activation loop S226 and Y229 to alanine. The

196   small-scale expression screen showed that the inactive mutants of PfGSK3 did not express in

197   a soluble form but instead formed inclusion bodies (Fig. 6A). We then isolated the inclusion

198   bodies of the PfGSK3^{K96A} and PfGSK3$^{S226A/Y229A}$ mutants and examined their phosphorylation

199   states by mass spectrometry. These mutants display the full length sequence but are devoid

200   of any phosphorylation (K96A) or exhibit markedly reduced phosphorylation (S226A/Y229A)

201   compared to the wild-type protein (Fig. 6C). This indicates that these mutations either impair

202   PfGSK3 function or PfGSK3 phosphorylation is crucial for folding and solubility. The latter is

203   further supported by a co-expression experiment of wild-type PfGSK3 with the

204   dephosphorylating lambda phosphatase, which results in insoluble material only (Fig. S6C).

205   As we found the N-terminal extension of PfGSK3 to carry several phosphorylation sites, we

206   cloned a series of N-terminally truncated protein constructs (Fig. S1A). None of these

207   constructs expressed in a soluble form (Fig. 6B) and mass spectrometry confirmed a complete

208   lack of phosphorylation in the PfGSK3 activation loop (ΔN64 in Fig. 6C). Expression tests at

209   different temperatures indicate that not only the aforementioned mutants and constructs, but

210   also wild-type PfGSK3 is less expressed and mainly insoluble at lower temperatures. In fact,

211   constructs with truncated N-termini did not express at 18°C at all, and even the wild-type

212   protein forms only inclusion bodies under such experimental settings (Fig. S6A-B). These data

213   further stress the crucial role of the PfGSK3 N-terminus as well as its phosphorylation sites for

214   folding and function of PfGSK3. As the N-terminally truncated PfGSK3 constructs are devoid

215   of any phosphorylation, we propose that the N-terminal PfGSK3 domain, unique for

216   *Plasmodium* species, is essential for PfGSK3 stability.

217

218   **N-terminal phosphorylation fine-tunes the activity of PfGSK3**

219   Provided that the PfGSK3 N-terminus is indispensable for its function, we suspected that this

220   region could also be a potential site of PfGSK3 regulation, as it is in human GSK3β. Therefore,

221   we analyzed the individual separated fractions from IEX chromatography (Fig. 1D) by

222   LC/MS-MS. Here, we observed that the extent of phosphorylation of residues in the activation

223   loop of PfGSK3 is comparable between the four fractions, but large differences in the

224   N-terminal phosphorylation patterns are obvious (Fig. 7A). Specifically, all N-terminal residues

225   are significantly less phosphorylated in the first fraction (F1), while in the second fraction (F2),

226   the extend of phosphorylation of residues Y39 and S40 is strongly increased and the third and

227   fourth fractions (F3 and F4) have generally higher level of N-terminal phosphorylation, which

228   is most pronounced on residues Y39, S40, S42 and S43. Of note, only residue S40 is well

229   conserved across *Plasmodium* species, but all other *Plasmodium* species contain multiple

230   non-conserved N-terminal residues that are potential phosphorylation sites (Fig. S7A, F). To

231   assess the relationship between the activity of PfGSK3 and its phosphorylation pattern, we

232   made use of the differences in N-terminal phosphorylation between the individual IEX fractions

233   and measured their kinase activity separately (Fig. 7B). PfGSK3 activity rises with an increase

234   in N-terminal phosphorylation levels (from F1 to F4), supporting that phosphorylation is crucial

235   for activity. Phosphorylation levels of residues S40, S42 and S43 significantly correlate with

236   the PfGSK3 activity (Pearson R=0.6, p<0.05 in all cases, Fig. 7C-E), while this correlation is

237   absent or insignificant for the other residues (Fig. S7B-F). These data support the hypothesis

238   that phosphorylation of N-terminal PfGSK3 residues enhances PfGSK3 activity.

239   **N-terminal phosphorylation affects the PfGSK3 structure**

240   The phosphorylated residues at the PfGSK3 N-terminus are, based on structural homologs

241   (24, 25), expected to be distant from the active site of the protein and it is therefore unclear

242   how N-terminal phosphorylation regulates the protein activity. In the human counterpart, the

243   phosphorylation of S9 prompts the binding of the N-terminus to a pocket for a priming

244   phosphorylation in the substrate-binding region and reduces the binding affinity for substrates

245   (24, 25). Hence, we investigated structural consequences of N-terminal phosphorylation in

246   PfGSK3. Circular dichroism data on the least (F1) and most (F4) phosphorylated fractions

247   show no pronounced differences in their secondary structure content (Fig. 8A). Thus, to

248   investigate the impact of the N-terminal phosphorylation on the tertiary structure of PfGSK3,

249   we used the *de novo* structure prediction server Robetta and the TrRefineRosetta method (70)

250　　to generate initial PfGSK3 models that were further fitted against the SAXS data measured on

251　　separated PfGSK3 IEX fractions F1 and F4 (Table 1, see Fig. 1D). Using the wild type PfGSK3

252　　sequence as well as PfGSK3 sequences with phosphomimetic mutations (see Methods), we

253　　generated 25 models (Fig. 8A) that display similar structure of the core domains but are

254　　structurally diverse in both termini (Fig. 8A). We identified a subset of models that provide the

255　　best fits to the experimental SAXS data (Table 1, Fig. S8B) of fractions F1 or F4 (Fig. 8D).

256　　Overall, the models with a larger radius of gyration (see the point sizes in Fig. 8D) display

257　　lower χ^2 (goodness-of-fit) values. Further refinement of these best-fitting extended models was

258　　conducted in the program CORAL (71) that enables assigning additional flexibility to existing

259　　models (Fig. 8E,H). This approach led to poorer fits when only C-terminus (residues 403-440)

260　　or both N-terminus and C-terminus (residues 1-64 and 403-440) were assigned flexibility

261　　(Fig. 8E,H). However, conducting refinement with a disordered C-terminus (residues 403-440)

262　　and a flexible stretch of residues between the N-terminus and the PfGSK3 core (residues

263　　47-63, conformation B, Fig. 8E,H) led to two refined models displaying lower χ^2 values

264　　compared to the SAXS data from fraction F4. Ultimately, for both fractions, the models with

265　　the lowest χ^2 values display α-helical structure at the N-terminus. This α-helix is extended in

266　　the less phosphorylated F1 fraction (Fig. 8B) and folds back in the more phosphorylated F4

267　　fraction (Fig. 8C). The raw SAXS data support such conformational differences, as both the

268　　calculated radius of gyration R_g and maximum intra-particle distance D_{max} decrease from F1

269　　to F4 fraction (R_g 3.4 vs 3.2 nm; D_{max} 13 vs 12 nm, respectively, see Table 1). The

270　　dimensionless Kratky plot data representation also suggests a more extended conformation

271　　for the F1 fraction (Fig. S8C), as the F1 curve broadens and shifts to higher sR_g values relative

272　　to F4. The refined models fit the SAXS data well (χ^2 values of 1.19 and 1.06, respectively, Fig.

273　　8F,G), but should be viewed cautiously as an average representation of the conformations

274　　present in solution that contains a mixture of differently phosphorylated species. In conclusion,

275　　the combinatorial approach indeed shows that upon phosphorylation, PfGSK3 becomes more

276　　flexible, enabling a conformational change in the N-terminus. This change presumably

277　　mediates the increased PfGSK3 activity upon phosphorylation.

11

Discussion

Identification and characterization of novel drug targets in malaria remain an important goal as drug-resistant parasites are a persisting threat in the fight against malaria. PfGSK3 has been identified as a potential drug target since its inhibitors also possess potent and selective anti-plasmodial activity. 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 is the main reason why we were not able to grow PfGSK3 crystals for detailed structural characterization (72). Moreover, the PfGSK3 C-terminus, predicted to be disordered, and its unusual N-terminus represent additional challenges in protein crystallization. As the expression of inactive PfGSK3 protein (K96A, S226A, Y229A) as well as the removal of heavily phosphorylated N-terminus result in insoluble protein that lacks phosphorylation, other protein constructs and strategies will likely need to be considered for the structure determination of 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) (37–39, 62–65). The additional phosphorylation sites at the N-terminus (Y39, S40, S42 and S43) might have occurred due to the heterologous expression, but could also hint towards additional phosphorylation sites present *in vivo* under certain conditions. In general, these phosphorylation sites are reminiscent of those identified in human GSK3β. In GSK3β, the autophosphorylation of Y216 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

12

305   autophosphorylation is likely conserved between the human and *P. falciparum* proteins, the

306   role of N-terminal phosphorylation seems to be different.

307   First, the N-terminal sequence is not conserved between the two homologous proteins, with

308   *Plasmodium* species displaying an unusual extension with a number of potential

309   phosphorylation sites. Second, whereas N-terminal phosphorylation of S9 in human GSK3 is

310   mediated by other kinases, the phosphorylation of N-terminal residues in PfGSK3 is the result

311   of autophosphorylation, as we show here. Last, while phosphorylation of S9 in human GSK3

312   inhibits its activity, our work illustrates that the increasing amount of phosphorylation on the

313   PfGSK3 N-terminus results in an increased enzymatic activity.

314   Our structural investigation using the combination of structure prediction and its validation

315   using experimental SAXS data provides further insights into how the N-terminal

316   phosphorylation affects PfGSK3 structure. Two SAXS datasets originating from two differently

317   phosphorylated IEX fractions of PfGSK3 directed the modelling towards PfGSK3

318   conformations with an α-helical N-terminus in both cases. However, this helix extends away

319   from the active side of PfGSK3 in the less phosphorylated fraction, or folds back towards the

320   center of the protein in the more phosphorylated fraction. It is possible that the "folded back"

321   conformation (Fig. 8C) could be mediated by an N-terminal intramolecular coiled-coil

322   interaction induced by phosphorylation, as the probability of coiled-coil formation increases

323   with the increasing number of N-terminal PfGSK3 phosphomimetic mutations *in silico* (Fig.

324   S8D). Interestingly, human GSK3β undergoes a similar conformational change, where upon

325   phosphorylation of S9, its N-terminus binds in its substrate-binding region and competitively

326   inhibits its activity (24, 25). However, N-terminal phosphorylation of PfGSK3 leads to an

327   increase in enzymatic activity, suggesting a different regulatory mechanism. The

328   conformational change upon phosphorylation might position the very N-terminus of PfGSK3

329   close to its core domains, potentially enabling intramolecular interactions that could enhance

330   its activity. Alternatively, the N-terminal extension might interact with the substrates to guide

331   them to the PfGSK3 active site. In both cases, the conformational change would be driven by

332   phosphorylation of residues 39-43, whereas the interacting residues are possibly confined to

333   the very N-terminus (residues 1-20) that is conserved in *Plasmodium* species (Fig. S6A,F).

334   It should be noted that the N-termini of PfGSK3 homologues are not well conserved even

335   among the representatives of *Apicomplexa*; however, a number of residues that can be

336   phosphorylated could still enable a similar mode of regulation as we have described here (Fig.

337   S7A,F). Indeed, S33 of *Plasmodium berghei* GSK3 (PBANKA_0410400), equivalent to S40 in

338   PfGSK3, was found to be phosphorylated in early gametogenesis (73). Residues equivalent

339   to S40 of PfGSK3 are conserved across all *Plasmodium* species and residues that can be

340   phosphorylated are also present in the vicinity of S40 in most *Plasmodium species* (equivalent

341   to residues 38, 39 and 42 in *P. falciparum*), suggesting that the mechanism of regulation *via*

342   N-terminal phosphorylation is preserved in all *Plasmodium* GSK3 proteins.

343   We have shown that this N-terminal phosphorylation, as well as the phosphorylation of the

344   activation loop, are the result of PfGSK3 exhibiting autophosphorylation. This result confirms

345   data in a recent publication, where the authors suspected that a high background in their

346   phosphorylation assay could be attributed to autophosphorylation (74). Autophosphorylation

347   of human GSK3β occurs directly after protein synthesis while the protein is associated with

348   chaperones, such as HSP90 (27). Interestingly, a major contaminant during our PfGSK3

349   purification process was *E. coli* DnaK that shares 48-57% identity with various *P. falciparum*

350   Hsp70 proteins, indicating that chaperone-assisted autophosphorylation might be conserved

351   across these taxa. Moreover, PfGSK3 was less soluble and had lower expression levels at

352   lower temperatures (25°C and 18°C), at which the chaperone levels in *E. coli* are also reduced

353   (75). This further supports that chaperons could be crucial interaction partner for the first

354   stages of PfGSK3 autophosphorylation, which are likely crucial for its solubility.

355   Additionally, we have revealed another means of PfGSK3 regulation: the divalent heavy metal

356   ions, such as zinc and copper, inhibit the activity of PfGSK3 by inducing the formation of

357   heterogeneous high-MW particles. The ionic radii of these cations are in a small range of

358   1.09-1.21 Å, indicating that the effect is mediated through a specific binding site. The fact that

359   the protein becomes more sensitive to the metal ions only after the cleavage of the N-terminal

360   affinity tag indicates that this potential binding site could be located in close proximity to the

361   N-terminus of PfGSK3. Indeed, a partially overlapping set of divalent cations inhibiting human

362   GSK3β can be found in the literature, including lithium, beryllium, zinc, copper and mercury

363   (17, 76, 77). Lithium and beryllium ions inhibit GSK3 by competing with magnesium ions that

364   mediate ATP binding (77–79), but to our knowledge, they have not been shown to induce the

365   formation of high-MW particles. To this date, there has been no structure of a GSK3 homolog

366   with a bound heavy metal ion deposited in the PDB database, and thus the inhibition

367   mechanism of heavy metal cations remains unknown. Importantly, as the high-MW particles

368   can dissociate back into active PfGSK3 monomers, it cannot be excluded that this type of

369   regulation is applied *in vivo*. The concentration of zinc in *Plasmodium* cytosol peaks at the late

370   blood stages and should be sufficient to induce the formation of high-MW PfGSK3 species

371   (80). Moreover, the concept of "auto-inhibitory polymerization" has already been described in

372   detail for human CK2 (81), further underlining the possible similarities between PfCK2 and

373   PfGSK3. It remains to be evaluated experimentally *in vivo* whether the PfGSK3 high-MW

374   particles play a role in its native environment.

375   In conclusion, our work provides new insights into the function of PfGSK3 and its regulation

376   by N-terminal phosphorylation and heavy metals, highlighting that phosphorylation influences

377   the structure of the PfGSK3 N-terminus that in turn results in increased kinase activity.

378   **Experimental procedures**

379   **Cloning and mutagenesis**

380   The vector with N-terminally His-tagged PfGSK3 was generated by PCR amplification of the

381   GSK3 coding sequence from *P. falciparum* cDNA followed by Ligation Independent Cloning

382   into HindIII/KpnI-cleaved plasmid pOPIN F (82) using the In-Fusion HD EcoDry Cloning Kit

383   (Takara Clontech) according to the manufacturer's instructions. The mutants S226A, Y229A

384   and S226A/Y229A were generated by overlap extension PCR amplification from the original

385   vector and Ligation Independent Cloning as described above. The wild-type protein and the

386   mutant K96A cloned in pET28a vector were ordered from GenScript. The N-terminally

387   truncated constructs were cloned by amplifying the sequence from the original vector and

388   subcloning into BsaI-cleaved plasmid pNIC28_Bsa4 by SLiCE cloning (83). The construct

389   pSLI-GSK3-GFP-glms was generated by PCR-amplification of 1013 bp of the C-terminal

390   GSK3 coding sequence from genomic DNA using primers GSK3_NotI_sense_SLI and

391   GSK3_AvrII_as_fored_int and cloning into pSLI-GFP-glms (84, 85) using NotI/AvrII restriction

392   sites.

393

394   **PfGSK3 expression and purification**

395   *E. coli* C41(DE3) culture transformed with the PfGSK3 coding vector was grown in TB medium

396   supplemented with 0.1 mg/ml ampicillin at 37°C, induced at $OD_{600nm}=0.7$ with 0.5 mM IPTG

397   and harvested after 4 hours. The pellets were resuspended in 5 mL of lysis buffer (20mM NaP

398   pH 7.5, 300mM NaCl, 15mM imidazole, 5% glycerol, 0.5mM TCEP, 1 mg/ml lysozyme, 5 U/ml

399   DNase, 1 Roche protease inhibitor tablet/100mL) per 1 g of wet weight culture. The

400   suspension was passed three times through an emulsifier at 15 000 psi, centrifuged at

401   19 000 x g and the supernatant was loaded on a pre-equilibrated His-Trap column. The

402   column was washed with lysis buffer supplemented with 40 mM imidazole and the protein was

403   eluted at increasing imidazole concentrations. The protein containing fractions were

404   concentrated (10 kDa cut-off concentrator) with the addition of 1 mg of 3C protease per 3 L of

16

405   culture volume and further separated on a Superdex 200 HiLoad column (GE Healthcare) in

406   50mM Tris pH 8.0, 20mM NaCl, 0.5mM TCEP. The protein containing fractions were

407   concentrated again, loaded on ResourceQ anion exchange column (GE Healthcare) and

408   eluted with increasing concentration of NaCl. Fractions representing different PfGSK3 species

409   were pooled together, concentrated, dialyzed against the analysis buffer (typically 20mM Tris

410   pH 8.0, 100mM NaCl, 0.5mM TCEP) and flash-frozen in liquid nitrogen and stored at -80°C.

411   All purification steps were performed at 4°C.

Expression and purification of insoluble PfGSK3 constructs

413   The insoluble PfGSK3 constructs (K96A, S226A/Y229A, ΔN64) were expressed as described

414   for the wild-type PfGSK3. The pellets of harvested bacteria were resuspended in 5 mL of lysis

415   buffer (20mM NaP pH 7.5, 300mM NaCl, 15mM imidazole, 5% glycerol, 0.5mM TCEP, 1

416   mg/ml lysozyme, 5 U/ml DNase, 1 Roche protease inhibitor tablet/100mL) per 1 g of wet

417   weight culture. The suspension was passed three times through an emulsifier at 15 000 psi,

418   centrifuged at 19 000 x g and the pellet was dissolved in a denaturing buffer (20mM NaP

419   (pH 7.5), 300mM NaCl, 25mM imidazole, 5% glycerol, 6M urea, 0.5mM TCEP) and incubated

420   with 1 mL NiNTA bead slurry (Sigma) for 1h, 4°C. The beads were subsequently washed twice

421   with 10 ml of the same buffer and subsequently eluted with 5 ml of the same buffer containing

422   extra 250 mM imidazole. The protein was concentrated (10 kDa c/o) and the sample for mass

423   spectrometry was separated by SDS-PAGE gel.

Small-scale expression tests

425   4 mL of TB supplemented with ampicillin was inoculated wit pre-cultured *E. coli* C41(DE3) that

426   had been transformed with the tested plasmids. The cultures were grown at 37°C until the

427   OD_{600nm} =0.6 and further either induced with 0.5mM IPTG and grown for 4 hours at 37°C or

428   25°C, or induced with 0.1mM IPTG and grown for 16 h at 18°C. OD_{600nm} was measured and

429   culture volume corresponding to OD_{600nm} of 2.0/mL was harvested. The pellets were

430   resuspended in 400 µl of lysis buffer (lysate fractions), then lysed by 10 min of vortexing with

431 100 µl glass beads and the cell debris were removed by centrifugation (soluble fractions). 5 µl

432 of each fraction was mixed with 10 µl of loading dye and run on SDS-PAGE in duplicates. One

433 gel of each set of samples was blotted on PVDF membrane (*BioRad*) The membranes were

434 blocked by 1% BSA in TBS-T, washed, incubated for 1h with a HisProbe-HRP conjugated

435 antibody (ThermoFisher), washed with TBS-T and the blot was developed using Super Signal

436 West Pico Substrate (ThermoFisher) and Super Signal West Femto Substrate (ThermoFisher)

437 in a 1:10 ratio.

Glutaraldehyde cross-linking

439 1 mg/ml PfGSK3 in PBS and 0.5mM TCEP was mixed with 8x molar excess of either EDTA

440 or $ZnCl_2$ and 0.01% glutaraldehyde. The reaction was quenched with 330mM Tris (final

441 concentration) and SDS-PAGE loading dye and SDS-PAGE was run for 150 minutes at 150V.

442 The gels were stained with Direct Blue stain.

Thermal shift assay

444 The stability of PfGSK3 under different buffer conditions and additives was measured by

445 nanoDSF (Prometheus NT.48, NanoTemper Technologies, GmbH). The protein

446 concentration was adjusted to 1 mg/ml. 10 µl of samples were loaded in the glass capillaries

447 and heated from 20°C to 90°C with a heating rate of 1°C/min. The fluorescence signals with

448 excitation wavelength of 280 nm and emission wavelengths of 330 and 350 nm were recorded

449 and the melting temperature was calculated as maximum of the derivative of the fluorescence

450 at 330 nm.

Circular dichroism

452 Circular dichroism was measured on a Chirascan CD spectrometer (Applied Photophysics).

453 The protein concentration was adjusted to 1 µM by 2 mM NaP (pH 7.5), 4 mM NaCl,

454 0.05 mM TCEP prior to the measurement. The circular dichroism spectrum was measured 10x

455 between 185 nm and 260 nm with 1 nm step in 1 mm quartz cuvette and analyzed by the

456 CDSSTR algorithm (86, 87) using DichroWeb (88).

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) or a Superose 6 Increase 10/300 column (100 µl sample). PBS supplemented with 0.5mM 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

0.5 mg/ml of PfGSK3 was incubated at 37°C with 2mM ATP or cAMP and 10mM $MgCl_2$ overnight. The samples were subsequently dialyzed at 4°C overnight in 50mM Tris pH 8.0, 20mM NaCl, 0.5mM 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 50mM Tris pH 8.0, 20mM NaCl, 0.5mM 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.

Small angle X-ray scattering

The SAXS data were collected at the P12 BioSAXS Beamline at the PETRA III storage ring (c/o DESY, Hamburg, Germany) (89). 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 (90) and further analyzed using the ATSAS suite (91) of programs to determine the molar mass, radius of gyration and real-space distance distribution. Scattering curves were computed from models using CRYSOL (92) with a constant subtraction and maximum order of harmonics set to 50. The rigid body modelling performed using CORAL (71), setting different

484   parts of models yielding best CRYSOL fits as disordered. These were residues 1-65 and

485   403-440 in conformation A, resides 47-65 and 403-440 in conformation B and residues

486   403-440 in conformation C. The raw data were uploaded to SASBDB under accession

487   numbers SASDL77 and SASDL87.

488   **Kinase assay**

489   To investigate enzymatic activity of recombinant PfGSK3 a commercial luminescence-based

490   kinase assay (KinaseGlo Plus, Promega) was used as previously described (93). Briefly, 20 ng

491   PfGSK3, 12 µM GS-1 peptide substrate (residues 636-661 of the human glycogen synthase

492   1; sequence YRRAAVPPSPSLSRHSSPHQ(pS)EDEEE; pS = prephosphorylated serine,

493   Promega) and 6 µM ATP (UltraPure, Promega) in the kinase reaction buffer (40 mM Tris-HCl

494   pH 7.5; 20 mM $MgCl_2$; 0.1 mg/mL BSA) were used as standard reaction mix (final reaction

495   volume 5 µl). The kinase reaction was performed for 30 min at 30°C followed by incubation

496   with KinaseGlo reagent according to the manufacturer's instructions. The reaction was

497   transferred to a solid white 384-well plate (NUNC, ThermoFisher) and the luminescence signal

498   was measured in an EnVision Multilabel Plate Reader (PerkinElmer, Integration time 0.5

499   sec/well). To investigate reaction kinetics, kinase and substrate concentrations were varied

500   between 0 and 100 ng or 0 and 15 µM respectively. To investigate autophosphorylation,

501   reactions were performed with kinase concentrations between 0 and 500 ng, 30 µM ATP and

502   without a substrate.

503   ***P. falciparum* culture**

504   Parasites were cultured in 5% 0^+ human erythrocytes (Blood bank, Universitätklinikum

505   Hamburg Eppendorf) in RPMI medium supplemented with 0.5% Albumax at 37°C in an

506   atmosphere of 1% O_2, 5% CO_2 and 94% N_2 as described previously (94). Parasitemia was

507   monitored by Giemsa-staining of methanol-fixed blood smears. The parasite strain 3D7 was

508   used in this study (95).

509   **Parasite imaging**

510   For live cell imaging parasite cultures were pelleted by centrifugation at 500x g for 1 min.

511   Sedimented cells were resuspended in an equal volume of the supernatant, 5 µl of the sample

512   were placed on a microscopy slide and covered with a cover slip. Fluorescence microscopy

513   images were taken with a Leica DM6 B microscope equipped with Leica DFC9000 GT camera

514   using an 100x immersion oil objective. Images were aquired with Leica Application Suite X

515   (LAS X) software and processed using Adobe Photoshop CS2. Nuclei were stained with 1

516   µg/ml DAPI (Biomol) in RPMI at 37°C for 15 min.

517   **Generation of transgenic parasite lines**

518   For Transfection, mature schizonts were enriched using 60% percoll and electroporated with

519   50 µg of plasmid DNA using a Lonza Nucleofector II (96). Transfectants were seleceted with

520   4nM WR99210 (Jacobus Pharmaceuticals). Stable integrant cell lines were generated as

521   previously described (84). Briefly, parasites carrying the WR99210-selected episomal plasmid

522   were cultivated in the presence of 400 µg/ml G418 (Sigma) to select for integrants with the

523   desired modification of the genomic locus. Correct genomic integration of the construct was

524   confirmed by PCR using genomic DNA from G418-selected integrants and 3D7 wildtype.

525   Genomic DNA was isolated using QIAamp DNA Mini Kit. For PCR analysis primers specific

526   for 3' and 5' integration as well as wildtype locus of PfGSK3 were used.

527   **Purification of GSK3-GFP from *P. falciparum***

528   For purification of GSK3-GFP, parasites were synchronized by treatment with 5% sorbitol for

529   10 min at 37°C and subsequently cultured in a volume of 50 ml with 5% hematocrit. Culture

530   medium was changed daily until parasitemia had reached 8-10%. Erythrocytes were lysed

531   with 0.03% saponin, isolated parasites were washed three times with ice cold PBS and

532   resuspended in 2 ml lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 1x

533   protease inhibitor mix, 50 U/mL DNase). The sample was triturated five times through a 27G-

534   needle (Braun) and further lysed by three cycles of freezing at -80°C followed by thawing on

535   ice. The lysates were supplemented with 1% DDM and 0.1% CHS and incubated for 3 hours

536   at 4°C. Subsequently, the samples were mixed with 30 μl of GFP-Trap bead slurry

537   (ChromoTek) and incubated on a rotatory wheel for 3 hours at 4°C. The beads were then

538   washed with 1 ml of 150 mM NaCl, 10 mM Tris, protease inhibitor mix (Roche), 0.1% DDM,

539   0.01% CHS pH 7.5 and with 1 ml of 300 mM NaCl, 10 mM Tris, protease inhibitor mix, 0.02%

540   DDM, 0.002% CHS pH 7.5. The trapped proteins were eluted with 30 μl of 6% SDS, 250 mM

541   Tris-HCl, pH 6.8, 71 mM TCEP at 95°C.

542   **Sample preparation for mass spectrometry**

543   In-solution samples from parasite cell cultures were processed as previously described (97).

544   To assess the phosphorylation status of the recombinant PfGSK3 (purified protein, separated

545   IEX fractions, insoluble constructs), each sample was separated on SDS-PAGE gel.

546   Coomassie-stained bands were excised, chopped into small pieces and transferred to 0.5 ml

547   Eppendorf tubes. For all following steps, buffers were exchanged by two consecutive 15 min

548   incubation steps of the gel pieces with 200 μl of acetonitrile (ACN), whereby the ACN was

549   removed after each step. Proteins were reduced by the addition of 200 μl of a 10 mM DTT

550   solution in 100 mM ammonium bicarbonate (AmBiC, Sigma Aldrich, A6141) and incubation at

551   56°C for 30 min. Proteins were alkylated by the addition of 200 μl of 55 mM chloroacetamide

552   (CAA)100 mM AmBiC solution and incubation for 20 min in the dark. A 0.1 μg/μl stock solution

553   of trypsin (Promega, V511A) in trypsin resuspension buffer (Promega, V542A) was diluted

554   with ice-cold 50 mM AmBiC buffer to achieve a final concentration of 1 ng/μl. 50 μl thereof

555   were added to gel pieces, which were incubated first for 30 min on ice and then over night at

556   37°C. Gel pieces were sonicated for 15 min, spun down and the supernatant was transferred

557   into a glass vial (VDS Optilab, 93908556). Remaining gel pieces were washed with 50 μl of

558   an aqueous solution of 50% ACN and 1% formic acid and sonicated for 15 min. The combined

559   supernatants were dried in a speedvac and reconstituted in 10 μl of an aqueous solution of

560   0.1% (v/v) formic acid.

561   **Mass spectrometry data acquisition**

562 The mass spectra of the initial phosphorylation site assessment of PfGSK3 (Fig. SB) was

563 performed as follows: An UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping

564 cartridge (Precolumn C18 PepMap100, 5 mm, 300 μm i.d., 5 μm, 100 Å) and an analytical

565 column (Acclaim PepMap 100. 75 × 50 cm C18, 3 mm, 100 Å) was used. Trapping was carried

566 out with a constant flow of trapping solution (0.05% trifluoroacetic acid in water) at 30 μL/min

567 onto the trapping column for 6 minutes. Subsequently, peptides were eluted via the analytical

568 column running solvent A (0.1% formic acid in water) with a constant flow of 0.3 μL/min, with

569 increasing percentage of solvent B (0.1% formic acid in acetonitrile) from 2% to 4% in 4 min,

570 from 4% to 8% in 2 min, then 8% to 28% for a further 37 min, in another 9 min. from 28%-

571 40%, and finally 40%-80% for 3 min followed by re-equilibration back to 2% B in 5 min. The

572 outlet of the analytical column was coupled directly to an Orbitrap QExactive™ plus Mass

573 Spectrometer (Thermo) using the Nanospray Flex™ ion source in positive ion mode. The

574 QExactive plus was operated in positive ion mode with a spray voltage of 2.2 kV and capillary

575 temperature of 275°C. Full scan MS spectra with a mass range of 350–1400 m/z were

576 acquired in profile mode using a resolution of 70,000 (maximum injections time of 100 ms,

577 AGC target was set to $3x10^6$ ions). Precursors were isolated using the quadrupole with a

578 window of 1.7 m/z and fragmentation was triggered by HCD in fixed collision energy mode

579 with fixed collision energy of 28%. MS2 spectra were acquired with the Orbitrap with a

580 resolution of 17.500 and a max injection time of 70 ms.

581 The mass spectra of all other experiments were recorded as follows: Peptides were analyzed

582 by LC-MS/MS on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scentific) as

583 previously described (98). To this end, peptides were separated using an Ultimate 3000 nano

584 RSLC system (Dionex) equipped with a trapping cartridge (Precolumn C18 PepMap100, 5

585 mm, 300 μm i.d., 5 μm, 100 Å) and an analytical column (Acclaim PepMap 100. 75 × 50 cm

586 C18, 3 mm, 100 Å) connected to a nanospray-Flex ion source. For the detection of

587 posttranslational modified peptides, peptides were loaded onto the trap column at 30 μl per

588 min using solvent A (0.1% formic acid) and peptides were eluted using a gradient from 2 to

23

589   80% Solvent B (0.1% formic acid in acetonitrile) over 60 min at 0.3 µl per min (all solvents

590   were of LC-MS grade). The Orbitrap Fusion Lumos was operated in positive ion mode with a

591   spray voltage of 2.4 kV and capillary temperature of 275 °C. Full scan MS spectra with a mass

592   range of 375–1200 m/z were acquired in profile mode using a resolution of 120,000 (maximum

593   injections time of 50 ms, AGC target was set to Standard with a max injection time of 50 ms.

594   Precursors were isolated using the quadrupole with a window of 1.2 m/z and fragmentation

595   was triggered by HCD in fixed collision energy mode with fixed collision energy of 34%. MS2

596   spectra were acquired with the Orbitrap with a resolution of 15.000 and a max injection time

597   of 54 ms.

598   **Mass spectrometry data analysis**

599   Acquired data were analyzed using IsobarQuant (99) and Mascot V2.4 (Matrix Science) using

600   a reverse UniProt FASTA database from *E. coli* (UP000000625) or *Plasmodium falciparum*

601   (UP000001450) including common contaminants. The following modifications were taken into

602   account: Carbamidomethyl (C, fixed), Acetyl (K, variable), Acetyl (Protein N-term, variable),

603   Oxidation (M, variable), as well as the variable modification Phospho (ST) and Phospho (Y).

604   The mass error tolerance for full scan MS spectra was set to 10 ppm and to 0.02 Da for MS/MS

605   spectra. A maximum of 2 missed cleavages were allowed. A minimum of 2 unique peptides

606   with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were

607   required on the peptide and protein level (100). Prior performing the search, the protein

608   sequences corresponding to the expressed PfGSK3 constructs were added to the database

609   for the expression host. The raw data have been uploaded to the PRIDE database with the

610   dataset identifier PXD026387.

611   **Correlation analysis**

612   Each biological replicate of mass spectrometry experiment was paired with the corresponding

613   biological replicate of the activity measurement. The relative phosphorylation of a residue was

614   calculated as a ratio of number of time the residue was identified as phosphorylated (Mascot

24

615 score >32 and MD score >=5) to number of times the residue was identified in the biological

616 sample. The relative activity was calculated as buffer-subtracted readout of the luciferase

617 assay relative to control. The strength of the linear association was calculated as Pearson

618 correlation coefficient using R (101).

Negative stain electron microscopy

620 4 µl of 0.01 mg/ml protein solution was applied to carbon-coated Cu/Pd grids and incubated

621 for 1 min. Excess liquid was removed with the filter paper (Whatman), and the grid was washed

622 twice with water for 5 s and 30 s. 4 µl of 2% uranyl acetate were applied twice for 5 s and 30 s.

623 The micrographs were collected on a Talos L120C (CSSB Cryo-EM multiuser facility) with

624 70 000x magnification.

Bioinformatics

626 The homologous protein sequences were aligned with the program MAFFT (102) and depicted

627 in Jalview. The secondary structure was predicted by Jpred (103). The PfGSK3 structures

628 were predicted using the Robetta server and the TrRefineRosetta method (70). To simulate

629 the different N-terminally phosphorylated species present in the protein fractions, we

630 generated five models for each of the following PfGSK3 sequences: wild-type sequence, three

631 sequences with a single phosphomimetic mutation (S40E, S42E, S43E) and one with a triple

632 mutation S40E/S42E/S43E (abbreviated SE3). Per-residue local error estimates of the models

633 calculated by Robetta (104) are written in the B-factor column of the PDB files. For

634 phylogenetic analysis, protein sequences with highest identity scores were identified for

635 individual organisms with the blastp suite of NCBI (105), the sequences were aligned with the

636 ClustalW method and the R package msa (106), and the phylogenetic tree was calculated in

637 PhyML with LG +G+I+F method (107). The phylogenetic tree with multiple sequence

638 alignment was plotted using R package ggtree (108).

639 **Data availability**

640 The mass spectrometry data are available in PRIDE database with dataset identifier

641 PXD026387. The SAXS data have been deposited in the SASBDB, accession numbers

642 SASDL77 and SASDL87. All other raw data are available on Figshare (doi:

643 10.6084/m9.figshare.14806887).

Supporting information

645 This article contains supporting information.

Acknowledgments

Conflict of interests

657 The authors declare that they have no conflicts of interest with the contents of this article.

References

1. World Health Organization (2019) *World Malaria Report. https://www.who.int/malaria*

2. Haldar, K., Bhattacharjee, S., and Safeukui, I. (2018) Drug resistance in Plasmodium. *Nat. Rev. Microbiol.* **16**, 156–170

3. Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, G., Karlsson, A., Al-lazikani, B., Hersey, A., Oprea, T. I., and Overington, J. P. (2017) A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34

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

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.* 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*. 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.* 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.* 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.* 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.* 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.* 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.* 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.* 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.* 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.* 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*. 10.1016/j.neurobiolaging.2010.01.006

17. Eldar-Finkelman, H., and Martinez, A. (2011) GSK-3 Inhibitors: Preclinical and Clinical

709   Focus on CNS. *Front. Mol. Neurosci.* **4**, 1–18

710   18.   Bourel, E., Grieco, S. F., and Jope, R. S. (2016) Glycogen synthase kinase-3 (GSK3):
711   regulation, actions, and diseases. 10.1016/j.pharmthera.2014.11.016.Glycogen

712   19.   Jope, R. S. (2011) Glycogen Synthase Kinase-3 in the Etiology and Treatment of Mood
713   Disorders. *Front. Mol. Neurosci.* 10.3389/fnmol.2011.00016

714   20.   Singh, K. K. (2013) An emerging role for Wnt and GSK3 signaling pathways in
715   schizophrenia. *Clin. Genet.* 10.1111/cge.12111

716   21.   De Sarno, P., Axtell, R. C., Raman, C., Roth, K. A., Alessi, D. R., and Jope, R. S. (2008)
717   Lithium Prevents and Ameliorates Experimental Autoimmune Encephalomyelitis. *J.*
718   *Immunol.* 10.4049/jimmunol.181.1.338

719   22.   Del Ser, T., Steinwachs, K. C., Gertz, H. J., Andrés, M. V., Gómez-Carrillo, B., Medina,
720   M., Vericat, J. A., Redondo, P., Fleet, D., and León, T. (2013) Treatment of Alzheimer's
721   disease with the GSK-3 inhibitor tideglusib: A pilot study. *J. Alzheimer's Dis.*
722   10.3233/JAD-2012-120805

723   23.   Lovestone, S., Boada, M., Dubois, B., Hüll, M., Rinne, J. O., Huppertz, H. J., Calero,
724   M., Andrés, M. V., Gómez-Carrillo, B., León, T., and Del Ser, T. (2015) A phase II trial
725   of tideglusib in alzheimer's disease. *J. Alzheimer's Dis.* 10.3233/JAD-141959

726   24.   Ter Haar, E., Coll, J. T., Austen, D. A., Hsiao, H. M., Swenson, L., and Jain, J. (2001)
727   Structure of GSK3β reveals a primed phosphorylation mechanism. *Nat. Struct. Biol.* **8**,
728   593–596

729   25.   Dajani, R., Fraser, E., Roe, S. M., Young, N., Good, V., Dale, T. C., and Pearl, L. H.
730   (2001) Crystal structure of glycogen synthase kinase 3β: Structural basis for phosphate-
731   primed substrate specificity and autoinhibition. *Cell.* **105**, 721–732

732   26.   Hughes, K., Nikolakaki, E., Plyte, S. E., Totty, N. F., and Woodgett, J. R. (1993)
733   Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation.
734   *EMBO J.* 10.1002/j.1460-2075.1993.tb05715.x

735   27.   Lochhead, P. A., Kinstrie, R., Sibbet, G., Rawjee, T., Morrice, N., and Cleghon, V.
736   (2006) A chaperone-dependent GSK3β transitional intermediate mediates activation-
737   loop autophosphorylation. *Mol. Cell.* **24**, 627–633

738   28.   Cole, A. R., Causeret, F., Yadirgi, G., Hastie, C. J., McLauchlan, H., McManus, E. J.,
739   Hernández, F., Eickholt, B. J., Nikolic, M., and Sutherland, C. (2006) Distinct priming
740   kinases contribute to differential regulation of collapsin response mediator proteins by
741   glycogen synthase kinase-3 in vivo. *J. Biol. Chem.* 10.1074/jbc.M513344200

742   29.   Cole, A. R., Knebel, A., Morrice, N. A., Robertson, L. A., Irving, A. J., Connolly, C. N.,
743   and Sutherland, C. (2004) GSK-3 phosphorylation of the Alzheimer epitope within
744   collapsin response mediator proteins regulates axon elongation in primary neurons. *J.*
745   *Biol. Chem.* 10.1074/jbc.C400412200

746   30.   Sutherland, C., Leighton, I. A., and Cohen, P. (1993) Inactivation of glycogen synthase
747   kinase-3β by phosphorylation: New kinase connections in insulin and growth-factor
748   signalling. *Biochem. J.* **2296**, 15–19

749   31.   Wichers, J. S., Scholz, J. A. M., Strauss, J., Witt, S., Lill, A., Ehnold, L. I., Neupert, N.,
750   Liffner, B., Lühken, R., Petter, M., Lorenzen, S., Wilson, D. W., Löw, C., Lavazec, C.,
751   Bruchhaus, I., Tannich, E., Gilberger, T. W., and Bachmann, A. (2019) Dissecting the
752   gene expression, localization, membrane topology, and function of the plasmodium
753   falciparum STEVOR protein family. *MBio.* 10.1128/mBio.01500-19

754   32.   Prinz, B., Harvey, K. L., Wilcke, L., Ruch, U., Engelberg, K., Biller, L., Lucet, I., Erkelenz,
755   S., Heincke, D., Spielmann, T., Doerig, C., Kunick, C., Crabb, B. S., Gilson, P. R., and
756   Gilberger, T. W. (2016) Hierarchical phosphorylation of apical membrane antigen 1 is
757   required for efficient red blood cell invasion by malaria parasites. *Sci. Rep.* **6**, 34479

758   33.   Mordue, D. G., Desai, N., Dustin, M., and Sibley, L. D. (1999) Invasion by Toxoplasma
759   gondii establishes a moving junction that selectively excludes host cell plasma
760   membrane proteins on the basis of their membrane anchoring. *J. Exp. Med.*

761   10.1084/jem.190.12.1783

762 34. Narum, D. L., and Thomas, A. W. (1994) Differential localization of full-length and
763   processed forms of PF83/AMA-1 an apical membrane antigen of Plasmodium
764   falciparum merozoites. *Mol. Biochem. Parasitol.* 10.1016/0166-6851(94)90096-5

765 35. Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S., and
766   Cowman, A. F. (2000) Apical membrane antigen 1 plays a central role in erythrocyte
767   invasion by Plasmodium species. *Mol. Microbiol.* 10.1046/j.1365-2958.2000.02175.x

768 36. Treeck, M., Zacherl, S., Herrmann, S., Cabrera, A., Kono, M., Struck, N. S., Engelberg,
769   K., Haase, S., Frischknecht, F., Miura, K., Spielmann, T., and Gilberger, T. W. (2009)
770   Functional analysis of the leading malaria vaccine candidate AMA-1 reveals an
771   essential role for the cytoplasmic domain in the invasion process. *PLoS Pathog.*
772   10.1371/journal.ppat.1000322

773 37. Treeck, M., Sanders, J. L., Elias, J. E., and Boothroyd, J. C. (2011) The
774   phosphoproteomes of plasmodium falciparum and toxoplasma gondii reveal unusual
775   adaptations within and beyond the parasites' boundaries. *Cell Host Microbe.* **10**, 410–
776   419

777 38. Solyakov, L., Halbert, J., Alam, M. M., Semblat, J.-P., Dorin-Semblat, D., Reininger, L.,
778   Bottrill, A. R., Mistry, S., Abdi, A., Fennell, C., Holland, Z., Demarta, C., Bouza, Y.,
779   Sicard, A., Nivez, M.-P., Eschenlauer, S., Lama, T., Thomas, D. C., Sharma, P.,
780   Agarwal, S., Kern, S., Pradel, G., Graciotti, M., Tobin, A. B., and Doerig, C. (2011)
781   Global kinomic and phospho-proteomic analyses of the human malaria parasite
782   Plasmodium falciparum. *Nat. Commun.* **2**, 565

783 39. Lasonder, E., Green, J. L., Camarda, G., Talabani, H., Holder, A. A., Langsley, G., and
784   Alano, P. (2012) The plasmodium falciparum schizont phosphoproteome reveals
785   extensive phosphatidylinositol and cAMP-protein kinase A signaling. *J. Proteome Res.*
786   **11**, 5323–5337

787 40. Vulliez-Le Normand, B., Tonkin, M. L., Lamarque, M. H., Langer, S., Hoos, S., Roques,
788   M., Saul, F. A., Faber, B. W., Bentley, G. A., Boulanger, M. J., and Lebrun, M. (2012)
789   Structural and functional insights into the malaria parasite moving junction complex.
790   *PLoS Pathog.* 10.1371/journal.ppat.1002755

791 41. Cao, J., Kaneko, O., Thongkukiatkul, A., Tachibana, M., Otsuki, H., Gao, Q., Tsuboi,
792   T., and Torii, M. (2009) Rhoptry neck protein RON2 forms a complex with microneme
793   protein AMA1 in Plasmodium falciparum merozoites. *Parasitol. Int.*
794   10.1016/j.parint.2008.09.005

795 42. Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J. F., and Lebrun, M. (2009) Export
796   of a Toxoplasma gondii rhoptry neck protein complex at the host cell membrane to form
797   the moving junction during invasion. *PLoS Pathog.* 10.1371/journal.ppat.1000309

798 43. Alexander, D. L., Mital, J., Ward, G. E., Bradley, P., and Boothroyd, J. C. (2005)
799   Identification of the moving junction complex of Toxoplasma gondii: A collaboration
800   between distinct secretory organelles. *PLoS Pathog.* 10.1371/journal.ppat.0010017

801 44. Alexander, D. L., Arastu-Kapur, S., Dubremetz, J. F., and Boothroyd, J. C. (2006)
802   Plasmodium falciparum AMA1 binds a rhoptry neck protein homologous to TgRON4, a
803   component of the moving junction in Toxoplasma gondii. *Eukaryot. Cell.*
804   10.1128/EC.00040-06

805 45. Leykauf, K., Treeck, M., Gilson, P. R., Nebl, T., Braulke, T., Cowman, A. F., Gilberger,
806   T. W., and Crabb, B. S. (2010) Protein kinase a dependent phosphorylation of apical
807   membrane antigen 1 plays an important role in erythrocyte invasion by the malaria
808   parasite. *PLoS Pathog.* 10.1371/journal.ppat.1000941

809 46. Hodder, A. N., Crewther, P. E., Matthewll, M. L. S. M., Reid, G. E., Moritz, R. L.,
810   Simpson, R. J., and Anders, R. F. (1996) The disulfide bond structure of Plasmodium
811   apical membrane antigen-1. *J. Biol. Chem.* 10.1074/jbc.271.46.29446

812 47. Patel, A., Perrin, A. J., Flynn, H. R., Bisson, C., Withers-Martinez, C., Treeck, M.,

813   Flueck, C., Nicastro, G., Martin, S. R., Ramos, A., Gilberger, T. W., Snijders, A. P.,
814   Blackman, M. J., and Baker, D. A. (2019) Cyclic AMP signalling controls key
815   components of malaria parasite host cell invasion machinery. *PLoS Biol.*
816   10.1371/journal.pbio.3000264

817   48.   Wilde, M. L., Triglia, T., Marapana, D., Thompson, J. K., Kouzmitchev, A. A., Bullen, H.
818   E., Gilson, P. R., Cowman, A. F., and Tonkina, C. J. (2019) Protein kinase A is essential
819   for invasion of Plasmodium falciparum into human erythrocytes. *MBio.* **10**, 1–15

820   49.   Fugel, W., Oberholzer, A. E., Gschloessl, B., Dzikowski, R., Pressburger, N., Preu, L.,
821   Pearl, L. H., Baratte, B., Ratin, M., Okun, I., Doerig, C., Kruggel, S., Lemcke, T., Meijer,
822   L., and Kunick, C. (2013) 3,6-Diamino-4-(2-halophenyl)-2-benzoylthieno[2,3-b]pyridine-
823   5-carbonitriles are selective inhibitors of Plasmodium falciparum glycogen synthase
824   kinase-3. *J Med Chem.* **56**, 264–275

825   50.   Masch, A., Nasereddin, A., Alder, A., Bird, M. J., Schweda, S. I., Preu, L., Doerig, C.,
826   Dzikowski, R., Gilberger, T. W., and Kunick, C. (2019) Structure-activity relationships
827   in a series of antiplasmodial thieno[2,3-b]pyridines. *Malar. J.* **18**, 1–10

828   51.   Moolman, C., van der Sluis, R., Beteck, R. M., and Legoabe, L. J. (2021) Exploration
829   of benzofuran-based compounds as potent and selective Plasmodium falciparum
830   glycogen synthase kinase-3 (PfGSK-3) inhibitors. *Bioorg. Chem.* **112**, 104839

831   52.   OSOLODKIN, D. I., ZAKHAREVICH, N. V., PALYULIN, V. A., DANILENKO, V. N., and
832   ZEFIROV, N. S. (2011) Bioinformatic analysis of glycogen synthase kinase 3: human
833   versus parasite kinases. *Parasitology.* **138**, 725–735

834   53.   Droucheau, E., Primot, A., Thomas, V., Mattei, D., Knockaert, M., Richardson, C.,
835   Sallicandro, P., Alano, P., Jafarshad, A., Baratte, B., Kunick, C., Parzy, D., Pearl, L.,
836   Doerig, C., and Meijer, L. (2004) Plasmodium falciparum glycogen synthase kinase-3:
837   Molecular model, expression, intracellular localisation and selective inhibitors. *Biochim.*
838   *Biophys. Acta - Proteins Proteomics.* **1697**, 181–196

839   54.   Dajani, R., Fraser, E., Roe, S. M., Yeo, M., Good, V. M., Thompson, V., Dale, T. C.,
840   and Pearl, L. H. (2003) Structural basis for recruitment of glycogen synthase kinase 3β
841   to the axin-APC scaffold complex. *EMBO J.* 10.1093/emboj/cdg068

842   55.   Henley, Z. A., Bax, B. D., Inglesby, L. M., Champigny, A., Gaines, S., Faulder, P., Le,
843   J., Thomas, D. A., Washio, Y., and Baldwin, I. R. (2017) From PIM1 to PI3Kδ via
844   GSK3β: Target Hopping through the Kinome. *ACS Med. Chem. Lett.*
845   10.1021/acsmedchemlett.7b00296

846   56.   Gentile, G., Merlo, G., Pozzan, A., Bernasconi, G., Bax, B., Bamborough, P., Bridges,
847   A., Carter, P., Neu, M., Yao, G., Brough, C., Cutler, G., Coffin, A., and Belyanskaya, S.
848   (2012) 5-Aryl-4-carboxamide-1,3-oxazoles: Potent and selective GSK-3 inhibitors.
849   *Bioorganic Med. Chem. Lett.* 10.1016/j.bmcl.2012.01.034

850   57.   Gentile, G., Bernasconi, G., Pozzan, A., Merlo, G., Marzorati, P., Bamborough, P., Bax,
851   B., Bridges, A., Brough, C., Carter, P., Cutler, G., Neu, M., and Takada, M. (2011)
852   Identification of 2-(4-pyridyl)thienopyridinones as GSK-3β inhibitors. *Bioorganic Med.*
853   *Chem. Lett.* 10.1016/j.bmcl.2011.06.050

854   58.   Chu, N., Salguero, A. L., Liu, A. Z., Chen, Z., Dempsey, D. R., Ficarro, S. B., Alexander,
855   W. M., Marto, J. A., Li, Y., Amzel, L. M., Gabelli, S. B., and Cole, P. A. (2018) Akt Kinase
856   Activation Mechanisms Revealed Using Protein Semisynthesis. *Cell.*
857   10.1016/j.cell.2018.07.003

858   59.   Bax, B., Carter, P. S., Lewis, C., Guy, A. R., Bridges, A., Tanner, R., Pettman, G.,
859   Mannix, C., Culbert, A. A., Brown, M. J. B., Smith, D. G., and Reith, A. D. (2001) The
860   structure of phosphorylated GSK-3β complexed with a peptide, FRATtide, that inhibits
861   β-catenin phosphorylation. *Structure.* 10.1016/S0969-2126(01)00679-7

862   60.   Liang, S. H., Chen, J. M., Normandin, M. D., Chang, J. S., Chang, G. C., Taylor, C. K.,
863   Trapa, P., Plummer, M. S., Para, K. S., Conn, E. L., Lopresti-Morrow, L., Lanyon, L. F.,
864   Cook, J. M., Richter, K. E. G., Nolan, C. E., Schachter, J. B., Janat, F., Che, Y.,
865   Shanmugasundaram, V., Lefker, B. A., Enerson, B. E., Livni, E., Wang, L., Guehl, N.

866   J., Patnaik, D., Wagner, F. F., Perlis, R., Holson, E. B., Haggarty, S. J., El Fakhri, G.,
867   Kurumbail, R. G., and Vasdev, N. (2016) Discovery of a Highly Selective Glycogen
868   Synthase Kinase-3 Inhibitor (PF-04802367) That Modulates Tau Phosphorylation in the
869   Brain: Translation for PET Neuroimaging. *Angew. Chemie - Int. Ed.*
870   10.1002/anie.201603797

871   61.   Wang, Y., Wach, J. Y., Sheehan, P., Zhong, C., Zhan, C., Harris, R., Almo, S. C.,
872   Bishop, J., Haggarty, S. J., Ramek, A., Berry, K. N., O'Herin, C., Koehler, A. N., Hung,
873   A. W., and Young, D. W. (2016) Diversity-Oriented Synthesis as a Strategy for
874   Fragment Evolution against GSK3β. *ACS Med. Chem. Lett.*
875   10.1021/acsmedchemlett.6b00230

876   62.   Pease, B. N., Huttlin, E. L., Jedrychowski, M. P., Talevich, E., Harmon, J., Dillman, T.,
877   Kannan, N., Doerig, C., Chakrabarti, R., Gygi, S. P., and Chakrabarti, D. (2013) Global
878   analysis of protein expression and phosphorylation of three stages of *Plasmodium*
879   *falciparum* intraerythrocytic development. *J. Proteome Res.* **12**, 4028–4045

880   63.   Ganter, M., Goldberg, J. M., Dvorin, J. D., Paulo, J. A., King, J. G., Tripathi, A. K., Paul,
881   A. S., Yang, J., Coppens, I., Jiang, R. H. Y., Elsworth, B., Baker, D. A., Dinglasan, R.
882   R., Gygi, S. P., and Duraisingh, M. T. (2017) Plasmodium falciparum CRK4 directs
883   continuous rounds of DNA replication during schizogony. *Nat. Microbiol.* **2**, 17017

884   64.   Kumar, S., Kumar, M., Ekka, R., Dvorin, J. D., Paul, A. S., Madugundu, A. K., Gilberger,
885   T., Gowda, H., Duraisingh, M. T., Keshava Prasad, T. S., and Sharma, P. (2017)
886   PfCDPK1 mediated signaling in erythrocytic stages of Plasmodium falciparum. *Nat.*
887   *Commun.* **8**, 1–12

888   65.   Flueck, C., Drought, L. G., Jones, A., Patel, A., Perrin, A. J., Walker, E. M., Nofal, S.
889   D., Snijders, A. P., Blackman, M. J., and Baker, D. A. (2019) Phosphodiesterase beta
890   is the master regulator of camp signalling during malaria parasite invasion. *PLoS Biol.*
891   **17**, e3000154

892   66.   Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J. V., Mijakovic, I., and Mann, M.
893   (2008) Phosphoproteome analysis of E. coli reveals evolutionary conservation of
894   bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteomics.* **7**, 299–307

895   67.   Deutscher, J., and Saier, M. H. (2006) Ser/Thr/Tyr protein phosphorylation in bacteria
896   - For long time neglected, now well established. *J. Mol. Microbiol. Biotechnol.* **9**, 125–
897   131

898   68.   Mann, N. H. (1984) Protein Phosphorylation in Escherichia coli and Purification of a
899   Protein Kinase. *J. Biol. Chem.* **259**, 526–533

900   69.   Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., and
901   Kuster, B. (2011) Confident phosphorylation site localization using the mascot delta
902   score. *Mol. Cell. Proteomics.* 10.1074/mcp.M110.003830

903   70.   Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., and Baker, D. (2020)
904   Improved protein structure prediction using predicted interresidue orientations. *Proc.*
905   *Natl. Acad. Sci. U. S. A.* 10.1073/pnas.1914677117

906   71.   Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M.,
907   Gorba, C., Mertens, H. D. T., Konarev, P. V., and Svergun, D. I. (2012) New
908   developments in the ATSAS program package for small-angle scattering data analysis.
909   *J. Appl. Crystallogr.* 10.1107/S0021889812007662

910   72.   Shrestha, A., Hamilton, G., O'Neill, E., Knapp, S., and Elkins, J. M. (2012) Analysis of
911   conditions affecting auto-phosphorylation of human kinases during expression in
912   bacteria. *Protein Expr. Purif.* **81**, 136–143

913   73.   Invergo, B. M., Brochet, M., Yu, L., Choudhary, J., Beltrao, P., and Billker, O. (2017)
914   Sub-minute Phosphoregulation of Cell Cycle Systems during Plasmodium Gamete
915   Formation. *Cell Rep.* **21**, 2017–2029

916   74.   Bui, H. T. N., Niederwieser, I., Bird, M. J., Dai, W., Brancucci, N. M. B., Moes, S., Jenoe,
917   P., Lucet, I. S., Doerig, C., and Voss, T. S. (2019) Mapping and functional analysis of

918   heterochromatin protein 1 phosphorylation in the malaria parasite Plasmodium
919   falciparum. *Sci. Rep.* 10.1038/s41598-019-53325-9

920   75.   Arsène, F., Tomoyasu, T., and Bukau, B. (2000) The heat shock response of
921   Escherichia coli. in *International Journal of Food Microbiology*, 10.1016/S0168-
922   1605(00)00206-3

923   76.   Ilouz, R., Kaidanovich, O., Gurwitz, D., and Eldar-Finkelman, H. (2002) Inhibition of
924   glycogen synthase kinase-3β by bivalent zinc ions: Insight into the insulin-mimetic
925   action of zinc. *Biochem. Biophys. Res. Commun.* 10.1016/S0006-291X(02)00636-8

926   77.   Ryves, W. J., Harwood, A. J., Dajani, R., and Pearl, L. (2002) Glycogen synthase
927   kinase-3 inhibition by lithium and beryllium suggests the presence of two magnesium
928   binding sites. *Biochem. Biophys. Res. Commun.* 10.1006/bbrc.2001.6305

929   78.   Ryves, W. J., and Harwood, A. J. (2001) Lithium inhibits glycogen synthase kinase-3
930   by competition for magnesium. *Biochem. Biophys. Res. Commun.*
931   10.1006/bbrc.2000.4169

932   79.   Mudireddy, S. R., Abdul, A. R. M., Gorjala, P., and Gary, R. K. (2014) Beryllium is an
933   inhibitor of cellular GSK-3β that is 1,000-fold more potent than lithium. *BioMetals*.
934   10.1007/s10534-014-9783-y

935   80.   Marvin, R. G., Wolford, J. L., Kidd, M. J., Murphy, S., Ward, J., Que, E. L., Mayer, M.
936   L., Pennen-Hahn, J. E., Haldar, K., and O'Halloran, T. V. (2012) Fluxes in 'Free' and
937   Total Zinc are Essential for Progression of Intraerythrocytic Stages of Plasmodium
938   falciparum. *Bone*. **19**, 731–741

939   81.   Lolli, G., Pinna, L. A., and Battistutta, R. (2012) Structural determinants of protein kinase
940   CK2 regulation by autoinhibitory polymerization. *ACS Chem. Biol.* **7**, 1158–1163

941   82.   Berrow, N. S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N.,
942   Stuart, D. I., and Owens, R. J. (2007) A versatile ligation-independent cloning method
943   suitable for high-throughput expression screening applications. *Nucleic Acids Res.*
944   10.1093/nar/gkm047

945   83.   Zhang, Y., Werling, U., and Edelmann, W. (2014) Seamless Ligation Cloning Extract
946   (SLiCE) cloning method. *Methods Mol. Biol.* **1116**, 235–244

947   84.   Birnbaum, J., Flemming, S., Reichard, N., Soares, A. B., Mesén-Ramírez, P., Jonscher,
948   E., Bergmann, B., and Spielmann, T. (2017) A genetic system to study Plasmodium
949   falciparum protein function. *Nat. Methods*. 10.1038/nmeth.4223

950   85.   Burda, P. C., Crosskey, T., Lauk, K., Zurborg, A., Söhnchen, C., Liffner, B., Wilcke, L.,
951   Pietsch, E., Strauss, J., Jeffries, C. M., Svergun, D. I., Wilson, D. W., Wilmanns, M.,
952   and Gilberger, T. W. (2020) Structure-Based Identification and Functional
953   Characterization of a Lipocalin in the Malaria Parasite Plasmodium falciparum. *Cell
954   Rep.* 10.1016/j.celrep.2020.107817

955   86.   Manavalan, P., and Johnson, W. C. (1987) Variable selection method improves the
956   prediction of protein secondary structure from circular dichroism spectra. *Anal.
957   Biochem.* 10.1016/0003-2697(87)90135-7

958   87.   Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from
959   circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods
960   with an expanded reference set. *Anal. Biochem.* 10.1006/abio.2000.4880

961   88.   Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from
962   circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*.
963   10.1002/bip.20853

964   89.   Blanchet, C. E., Zozulya, A. V., Kikhney, A. G., Franke, D., Konarev, P. V., Shang, W.,
965   Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., Svergun, D. I., and Roessle, M.
966   (2012) Instrumental setup for high-throughput small-and wide-angle solution scattering
967   at the X33 beamline of EMBL Hamburg. *J. Appl. Crystallogr.*
968   10.1107/S0021889812013490

969   90.   Graewert, M. A., Franke, D., Jeffries, C. M., Blanchet, C. E., Ruskule, D., Kuhle, K.,

Flieger, A., Schäfer, B., Tartsch, B., Meijers, R., and Svergun, D. I. (2015) Automated pipeline for purification, biophysical and X-ray analysis of biomacromolecular solutions. *Sci. Rep.* 10.1038/srep10734

91.   Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M., and Svergun, D. I. (2017) ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* 10.1107/S1600576717007786

92.   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.* 10.1107/S0021889895007047

93.   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*

94.   Trager, W., and Jensen, J. B. (1976) Human malaria parasites in continuous culture. *Science (80-.).* 10.1126/science.781840

95.   Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, W. T., Corcoran, L. M., Burkot, T. R., and Carter, R. (1987) Genetic analysis of the human malaria parasite Plasmodium falciparum. *Science (80-.).* 10.1126/science.3299700

96.   Moon, R. W., Hall, J., Rangkuti, F., Ho, Y. S., Almond, N., Mitchell, G. H., Pain, A., Holder, A. A., and Blackman, M. J. (2013) Adaptation of the genetically tractable malaria pathogen Plasmodium knowlesi to continuous culture in human erythrocytes. *Proc. Natl. Acad. Sci. U. S. A.* 10.1073/pnas.1216457110

97.   Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., and Krijgsveld, J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* 10.15252/msb.20145625

98.   Sridharan, S., Kurzawa, N., Werner, T., Günthner, I., Helm, D., Huber, W., Bantscheff, M., and Savitski, M. M. (2019) Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat. Commun.* 10.1038/s41467-019-09107-y

99.   Franken, H., Mathieson, T., Childs, D., Sweetman, G. M. A., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G., Reinhard, F. B. M., Huber, W., and Savitski, M. M. (2015) Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* 10.1038/nprot.2015.101

100. 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.* 10.1074/mcp.M114.046995

101. R Core Team (2020) (2020) R: A language and environment for statistical computing. *R A Lang. Environ. Stat. Comput. R Found. Stat. Comput. Vienna, Austria*

102. 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.* 10.1093/nar/gkz342

103. Cole, C., Barber, J. D., and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* 10.1093/nar/gkn238

104. 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.* 10.1038/s41467-021-21511-x

105. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009) BLAST+: Architecture and applications. *BMC Bioinformatics.* **10**, 421

106. Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015) Msa: An R package for multiple sequence alignment. *Bioinformatics.* **31**, 3997–3999

107. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies:

1022   Assessing the performance of PhyML 3.0. *Syst. Biol.* 10.1093/sysbio/syq010

1023 108. Yu, G. (2020) Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc.*
1024   *Bioinforma.* **35**, 3041–3043

1025 109. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: A protein
1026   secondary structure prediction server. *Nucleic Acids Res.* 10.1093/nar/gkv332

1027 110. Prommana, P., Uthaipibull, C., Wongsombat, C., Kamchonwongpaisan, S., Yuthavong,
1028   Y., Knuepfer, E., Holder, A. A., and Shaw, P. J. (2013) Inducible Knockdown of
1029   Plasmodium Gene Expression Using the glmS Ribozyme. *PLoS One.* **8**, e73783

1030 111. Daubenberger, C. A., Tisdale, E. J., Curcic, M., Diaz, D., Silvie, O., Mazier, D., Eling,
1031   W., Bohrmann, B., Matile, H., and Pluschke, G. (2003) The N'-terminal domain of
1032   glyceraldehyde-3-phosphate dehydrogenase of the apicomplexan Plasmodium
1033   falciparum mediates GTPase Rab2-dependent recruitment to membranes. *Biol. Chem.*
1034   **384**, 1227–1237

1035

1036 **Abbreviations**

1037 AMA1   *Plasmodium falciparum* Apical membrane Antigen 1

1038 GSK3β   Glycogen Synthase Kinase beta

1039 IEX   ion exchange chromatography

1040 IMAC   immobilized metal affinity chromatography

1041 high-MW   high molecular weight

1042 PfGSK3   *Plasmodium falciparum* Glycogen Synthase Kinase

1043 SEC   size exclusion chromatography

1044

1045 **Table 1.** SAXS sample details, data acquisition parameters, structural parameters and

1046 atomistic modelling.

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})	48250	
Molecular weight from chemical composition (Da)	52045.2	
Concentration (analysis or injection, mg/ml)	10	
Solvent composition	20mM Tris pH 8.0, 100mM NaCl, 0.5mM 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 x 1 s)	900 s (900 x 1 s)
Mode	SEC-SAXS	SEC-SAXS
Sample temperature (°C)	20	20
Structural parameters		

35

Guinier analysis		
$I(0)$ (cm^{-1})	0.037±0.001	0.021±0.001
R_g (Å)	34.3±0.4	31.5±0.2
q-range (Å^{-1})	0.01-0.04	0.01-0.04
Fidelity (Quality of fit parameter, *AutoRg*)	0.58	0.98
Molecular weight (kDa)*		
From V_c	66.2	48.7
From *MoW*	86.0	34.1
P(r) analysis		
$I(0)$ (cm^{-1})	0.037±0.001	0.021±0.001
R_g (Å)	35.8±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	0.62	0.82
Porod volume x 10^3 (Å^3)	105	101
Atomistic modelling**	*CRYSOL* with constant subtraction and maximum order of harmonics equal to 50	*CORAL* hybrid rigid body modeling
Structures	Robetta model WLT (4)	
q-range for modelling (Å^{-1})	0.01-0.50	0.01-0.58
χ^2	1.19	1.06
Predicted R_g (Å)	35.3	
Starting crystal structures		Robetta model 3SE (4)
Flexible residues		47-63, 403-440
SASBDB IDs for data and models		
	SASDL77	SASDL87

1047 * calculated using the appropriate functions in ATSAS (91).

1048 ** The values for the best model for each respective dataset are shown.

1049   Figure 1

1050
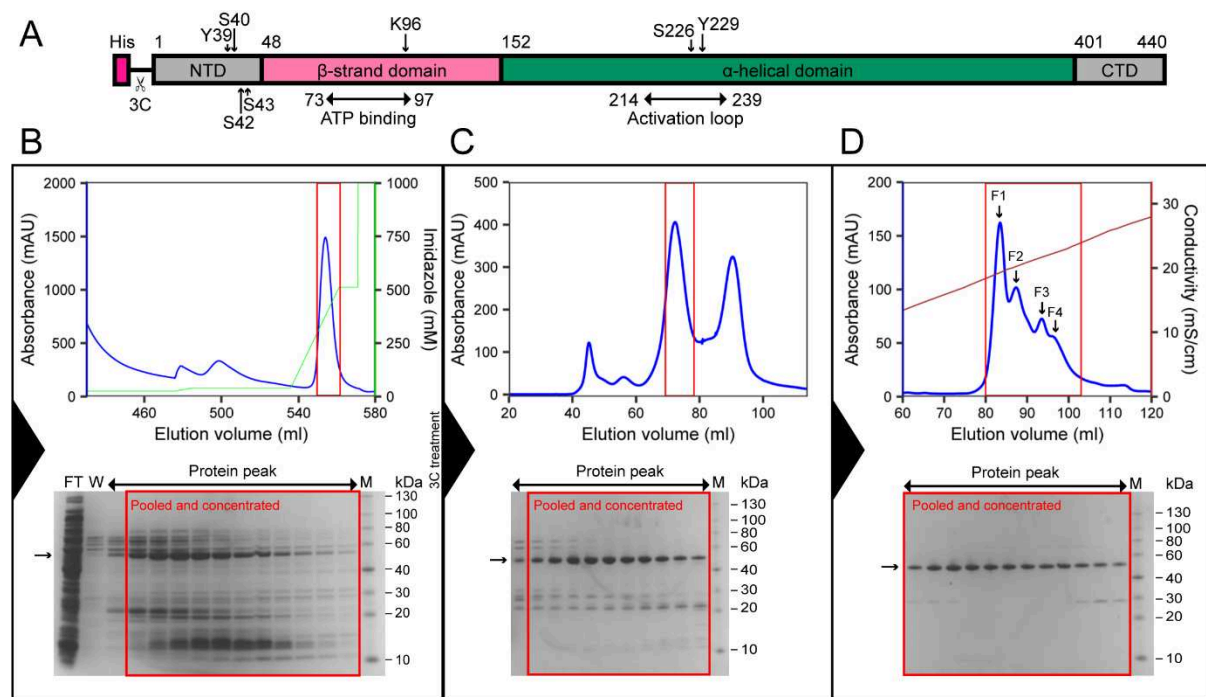
1051   **Fig. 1. Expression and purification of PfGSK3.**

1052   (A) Construct of PfGSK3 used for expression. The construct consists of the full-length

1053   sequence of PfGSK3 with an N-terminal His-tag and a 3C cleavage site. The domain

1054   organization and phosphorylation sites are marked. NTD is the N-terminal domain, CTD is the

1055   C-terminal domain, 3C is the 3C protease cleavage cite. (B) First PfGSK3 purification step:

1056   imidazole gradient elution profile from the His-Trap column. (C) Elution profile from the

1057   Superdex 200 size exclusion column. (D) NaCl gradient elution profile from Resource Q ion

1058   exchange column. The peaks of the ion exchange elution represent fractions that are

1059   phosphorylated to different extent, increasing from F1 to F4. The blue curves in the

1060   chromatograms show UV absorbance, the green curve shows imidazole concentration and

1061   the brown curve shows the conductivity. The peaks in red squares in chromatograms

1062   correspond to the red squares in corresponding SDS-PAGE gels.

37

1063   Figure 2

1064
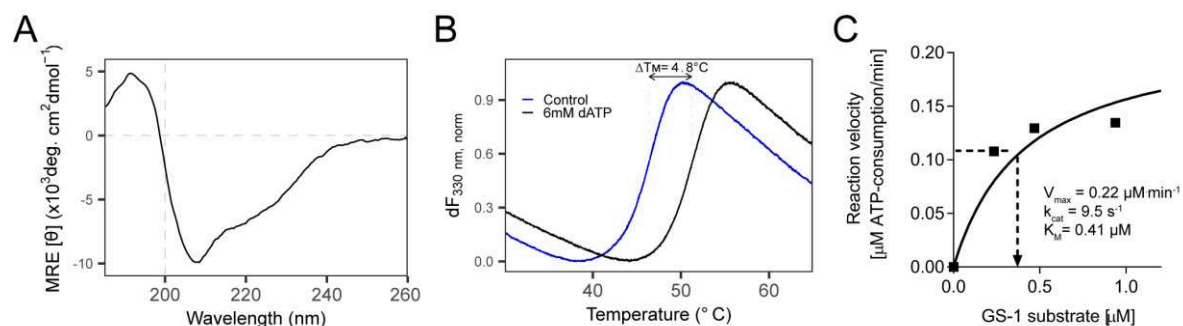
1065   **Fig. 2. Characterization of PfGSK3.**

1066   (A) Circular dichroism shows secondary structure composition similar to what is expected for

1067   a GSK3 protein, with 22% α helix, 26% β sheet, 24% turns and 29% disordered content. The

1068   circular dichroism was measured 10× and the data were averaged, buffer subtracted, and

1069   analyzed by DichroWeb. (B) Thermal unfolding profiles of PfGSK3 in presence or absence of

1070   6mM dATP. dATP stabilizes PfGSK3, suggesting that it binds in the ATP binding pocket. (C)

1071   Catalytic activity of GSK3 follows Michaelis-Menten kinetics. Saturation curve shows the

1072   reaction velocity (as calculated in FigS2F) plotted against different substrate concentrations.

1073   Michaelis-Menten constant (K_m), maximal reaction velocity (V_{max}) and turnover number (k_{cat})

1074   were calculated in GraphPad Prism using nonlinear regression. GS-1 = substrate peptide of

1075   the human glycogen synthase.

1076   Figure 3

1077
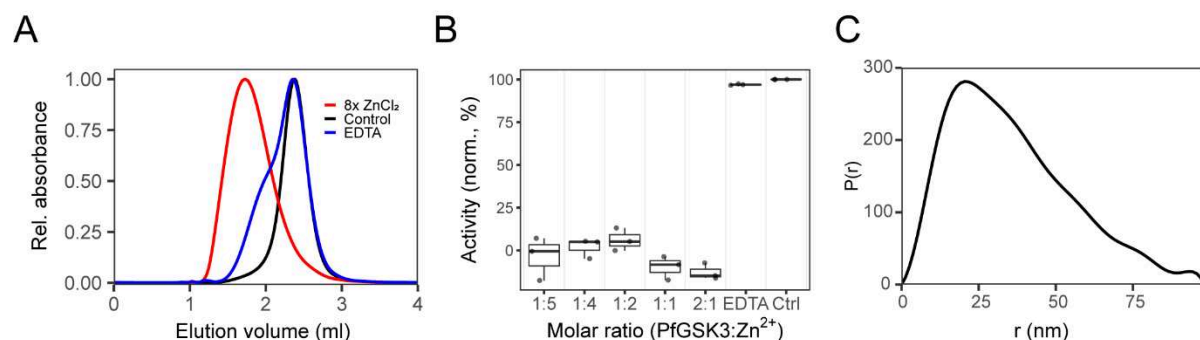
1078   **Fig. 3. Heavy metal ions inhibit PfGSK3 activity.**

1079   (A) Analytical size exclusion chromatography profiles on a Superose 6 column of PfGSK3

1080   apoprotein (black), after addition of 8× molar excess of zinc chloride (red) and after

1081   subsequent addition of EDTA (blue). The analysis shows that zinc ions induce the formation

1082   of high-MW PfGSK3 species. The formation is reversible because addition of EDTA shifts the

1083   elution profiles towards higher elution volumes. (B) The activity of PfGSK3 is strongly reduced

1084   in the presence of zinc ions, whereas the protein regains the activity after the addition of the

1085   metal chelator EDTA. (C) The distance distribution of the PfGSK3 high-MW species induced

1086   by zinc ions derived from SAXS data shows high heterogeneity in the sample, peaking at

1087   20 nm but with D_{max}=89 nm.

1088   Figure 4
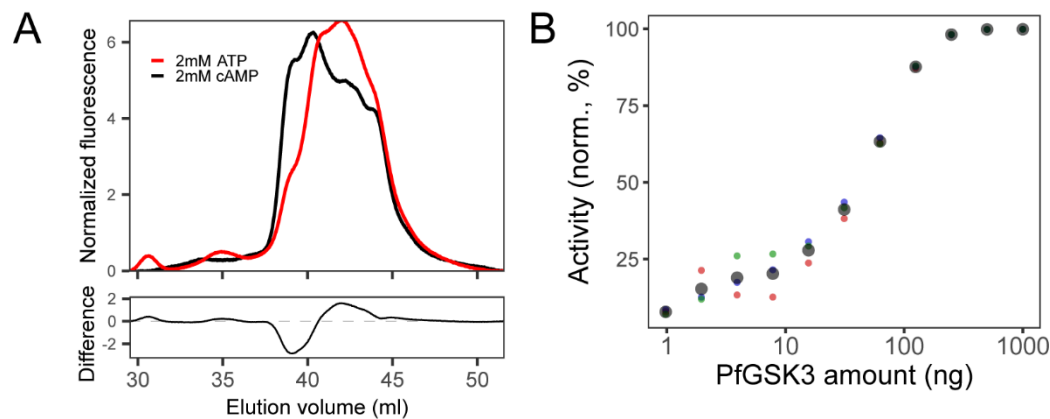
1089

1090   **Fig. 4. PfGSK3 exhibits autophosphorylation.**

1091   (A) IEX elution chromatograms (upper plot) of PfGSK3 after incubation with ATP (red) or cAMP

1092   as negative control (black) in the presence of magnesium ions suggest that PfGSK3 exhibits

1093   autophosphorylation. The difference between both chromatograms (bottom plot) shows an

1094   increase of later-eluting species after ATP treatment, indicating higher amount of

1095   phosphorylation. (B) The ATP consumption measured by the luminescence assay in the

1096   absence of the substrate is dependent on the enzyme concentration, which is indicative of the

1097   autophosphorylation exhibited by PfGSK3.

1098   **Figure 5**

1099
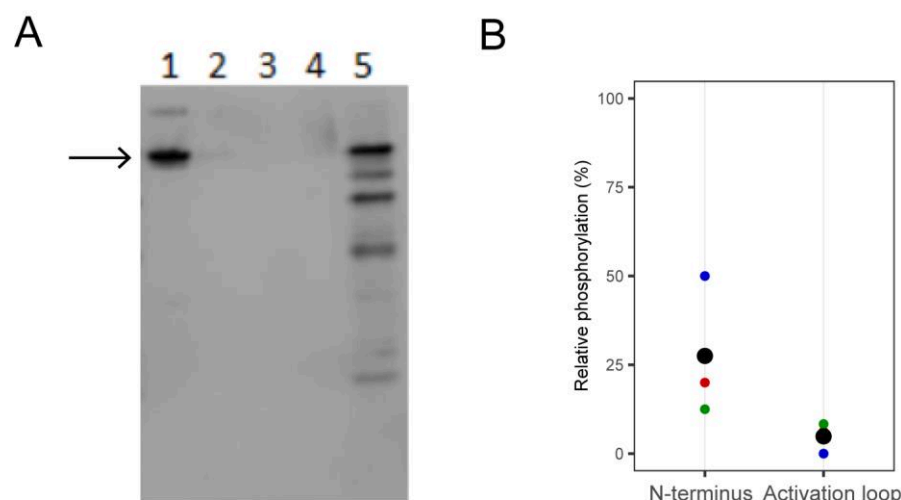
1100   **Fig. 5. N-terminus of PfGSK3 is phosphorylated *in vivo*.**

1101   (A) Western blot with anti-GFP antibodies shows the purification of PfGSK3-GFP from the

1102   plasmodium lysate (1), flow-through (2), two wash steps (3-4) and elution fraction (5). (B) The

1103   mass spectrometry analysis revealed that the N-terminal residues contain higher relative

1104   amount of phosphorylation (27%) compared to the activation loop (5%). The colored points

1105   are ratios from individual mass spectrometry experiment replicates and the large black points

1106   are their averages.

1107   Figure 6

1108
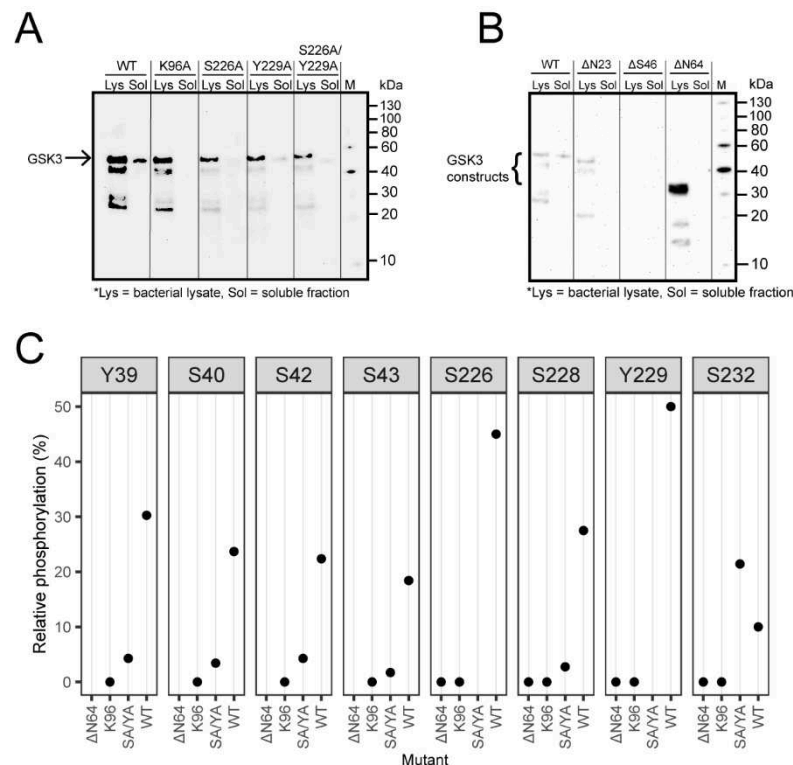
Fig. 6. The N-terminus and phosphorylation is essential for PfGSK3

1109

1110   (A) Small-scale expression test of PfGSK3 mutants with inactivating mutations in the ATP

1111   binding site (K96A) and activation loop (Y226A and S229A). Whereas all protein constructs

1112   are expressed, only the wild type PfGSK3 is soluble, which indicates that the

1113   autophosphorylation of PfGSK3 is important for its solubility. (B) Small-scale expression test

1114   of PfGSK3 N-truncated constructs that start with residues N23, S46 or N64. In spite of a very

1115   strong expression of PfGSK3-N64, none of the proteins is soluble, indicating that the PfGSK3

1116   N-terminus is crucial for the autophosphorylation process. The bacterial lysates (Lys) and their

1117   soluble fractions (Sol) were analyzed by Western blot with anti-His antibodies. (C) Analysis of

1118   phosphorylation of PfGSK3 mutants shows that the phosphorylation is completely lost (K96A,

1119   ΔN64) or reduced (S226A/Y229A) compared to the wild-type protein (WT).

1120   Figure 7

1121
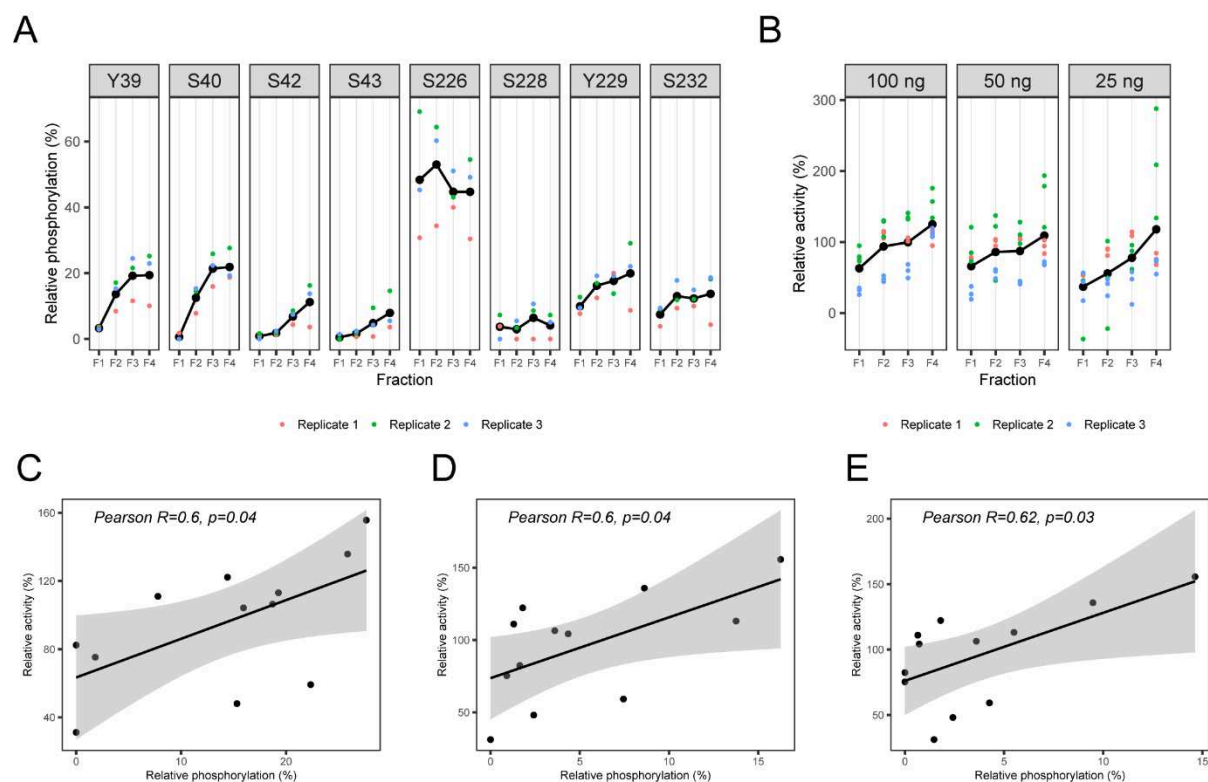
Fig. 7. N-terminal phosphorylation promotes PfGSK3 activity.

1123   (A) Relative phosphorylation of selected residues in the individual PfGSK3 fractions separated

1124   by ion exchange chromatography. The relative phosphorylation represents the fraction of all

1125   identified residues that were also identified as phosphorylated by LC-MS/MS with Mascot

1126   score > 32 and MD score ≥ 5. The data show that the residues in the N-terminal domain are

1127   gradually more phosphorylated with increasing IEX elution volume, whereas the

1128   phosphorylation at the activation loop remains constant. The different colors represent three

1129   biological triplicates. (B) Activity of individual PfGSK3 fractions separated by ion exchange

1130   chromatography measured with different PfGSK3 amount (100, 50 and 25 ng). The data show

1131   a general trend towards higher activity with increasing phosphorylation in the N-terminal

1132   domain. The activity was measured in biological triplicates and at different PfGSK3

1133   concentrations. (C-E) Correlation between the relative phosphorylation of the residues S40,

1134   S42 and S43 measured by mass spectrometry, respectively, and the relative activity of the

43

1135   same samples (triplicates of fractions F1-F4). The data were fitted with a linear model (black

1136   line), the gray background shows the 95% confidence interval.

1137   Figure 8

1138
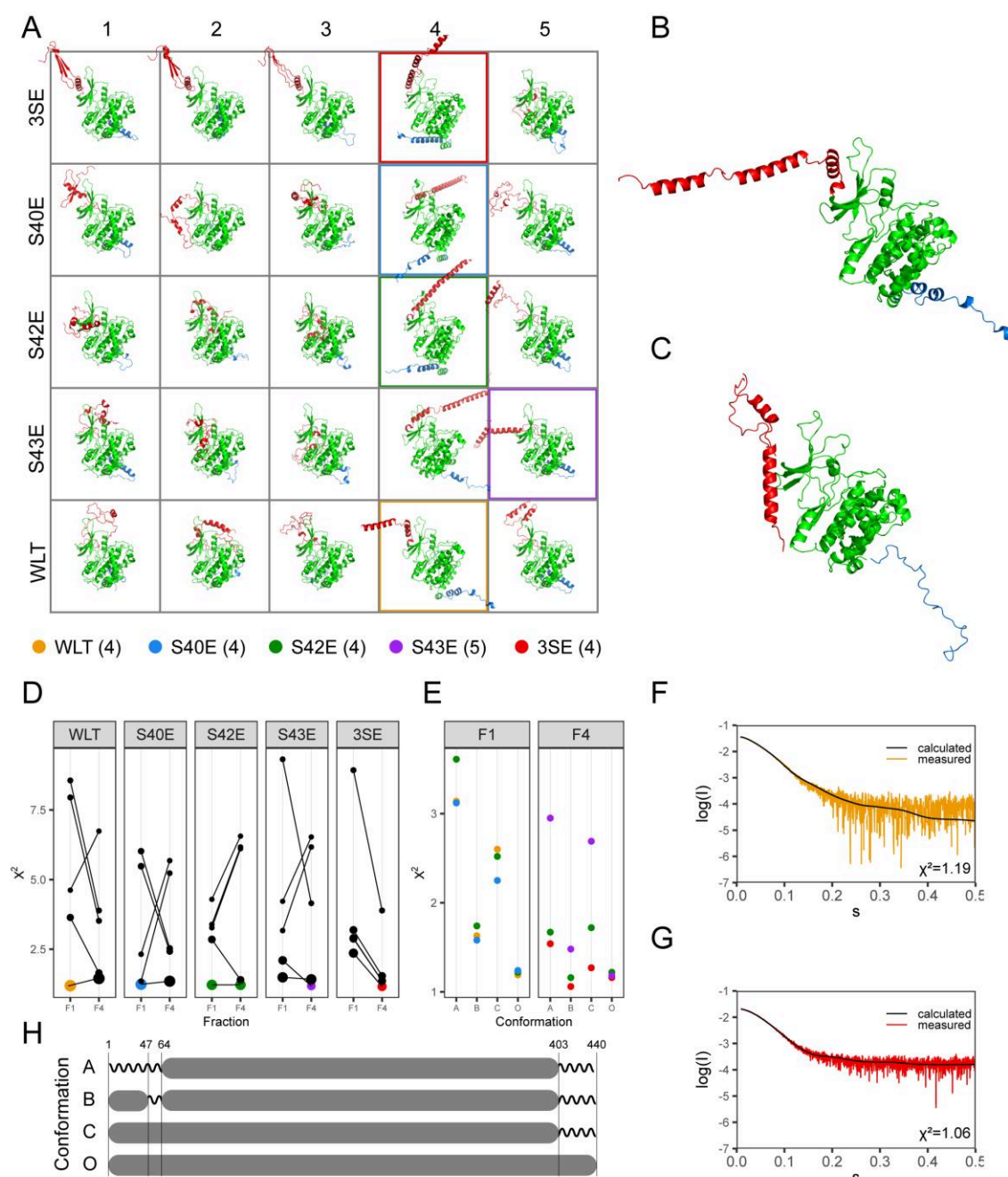
1139   **Fig. 8. The N-terminus of PfGSK3 changes the structure upon phosphorylation**

1140   (A) Structural models of PfGSK3 predicted by the Robetta server. Five different sequences

1141   were used as an input for the prediction and five models were predicted for each sequence.

1142   N-termini (residues 1-65) are red, the core domains green and C-termini (residues 403-440)

1143   blue. (B) The Robetta model that fits the SAXS data from F1 fraction best display an extended

1144   N-terminal helix. (C) The Robetta model refined against SAXS data from fraction F4 using

45

1145 CORAL with the best fit to the data. The N-terminal helix of this model folds back towards the

1146 core of PfGSK3. (D) X^2 values of PfGSK3 structures modelled based on five different

1147 sequence modifications (WLT, S40E, S42E, S43E and 3SE) compared to SAXS data recorded

1148 for fractions F1 or F4. The identical models are connected with lines. The size of each points

1149 correlates with the radius of gyration of the particular model. Three models with lowest X^2

1150 values for each F1 and F4 are colored (see legend). (E) X^2 values of the chosen models

1151 refined with CORAL using different strategies, enabling flexibility at either the N-terminus

1152 (residues 1-63, strategy A), C-terminus (residues 403-440, strategy C), or at both the C-

1153 terminus and residues 47-63 (strategy B). X^2 values of the original models are shown under

1154 "O". (F) The SAXS data measured on the PfGSK3 fraction F1 with the fit of theoretical

1155 scattering calculated from the best fitting model (X^2=1.19). (G) The SAXS data measured on

1156 the PfGSK3 fraction F1 with the fit of theoretical scattering calculated from the best fitting

1157 CORAL-refined model (X^2=1.06). (H) Graphical summary of the strategies used for model

1158 refinement against SAXS data using CORAL. The gray box symbolizes that the structure of

1159 the model was maintained and the zigzag line symbolizes that the structure was replaced with

1160 disordered residues by CORAL.

46