

SCIENTIFIC REPORTS



OPEN

Structure of ThiM from Vitamin B1 biosynthetic pathway of *Staphylococcus aureus* – Insights into a novel pro-drug approach addressing MRSA infections

Julia Drebes^{1,2,*}, Madeleine Künz^{1,*}, Björn Windshügel³, Alexey G. Kikhney⁴, Ingrid B. Müller², Raphael J. Eberle^{1,†}, Dominik Oberthür^{1,‡}, Huaixing Cang⁵, Dmitri I. Svergun⁴, Markus Perbandt^{1,7}, Christian Betzel^{1,7} & Carsten Wrenger⁶

Infections caused by the methicillin-resistant *Staphylococcus aureus* (MRSA) are today known to be a substantial threat for global health. Emerging multi-drug resistant bacteria have created a substantial need to identify and discover new drug targets and to develop novel strategies to treat bacterial infections. A promising and so far untapped antibiotic target is the biosynthesis of vitamin B1 (thiamin). Thiamin in its activated form, thiamin pyrophosphate, is an essential co-factor for all organisms. Therefore, thiamin analogous compounds, when introduced into the vitamin B1 biosynthetic pathway and further converted into non-functional co-factors by the bacterium can function as pro-drugs which thus block various co-factor dependent pathways. We characterized one of the key enzymes within the *S. aureus* vitamin B1 biosynthetic pathway, 5-(hydroxyethyl)-4-methylthiazole kinase (SaThiM; EC 2.7.1.50), a potential target for pro-drug compounds and analyzed the native structure of SaThiM and complexes with the natural substrate 5-(hydroxyethyl)-4-methylthiazole (THZ) and two selected substrate analogues.

The ability of bacteria to adapt to antibiotic exposure led to the continuous development and spread of multi-drug resistant microbes causing severe infections and even death¹. One of these pathogens is *Staphylococcus aureus* (*Sa*), which is currently resistant against a broad range of the most useful antibiotics^{2,3}. Despite the emergence of antibiotic resistance and relatively high mortality rates, the development of anti-infectives has seemingly been ignored for decades as a focus of drug discovery research investigations⁴⁻⁶. Most classical antibiotics inhibit cell wall linkage, interfere with the protein biosynthesis or decline nucleic acid synthesis. *Sa* developed strategies to obviate these impaired processes⁷. According to the WHO the 21st century could become the post-antibiotic era and the World Health Assembly commissioned a global action plan in May 2014⁸. These and other worldwide governmental activities clearly highlight the urgent need for incentives to develop agents that can overcome acquired resistance to treat bacterial and in particular staphylococcal infections.

¹University Hamburg, c/o DESY, Laboratory for Structural Biology of Infection and Inflammation, Hamburg, Germany. ²Department of Biochemistry, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany. ³Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Hamburg, Germany. ⁴EMBL Hamburg, c/o DESY, Hamburg, Germany. ⁵School of Life Sciences North Western Polytechnical University, Xi'an, Shaanxi, China. ⁶Unit for Drug Discovery, Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil. ⁷The Hamburg Centre for Ultrafast Imaging, Luruper Chaussee 149, D-22761 Hamburg, Germany. [†]Present address: Multiuser Center for Biomolecular Innovation, Department of Physics, Universidade Estadual Paulista (UNESP), São Jose do Rio Preto-SP 15054-000, Brazil. [‡]Present address: Center for Free-Electron Laser Science, Deutsches Elektronen Synchrotron-DESY, Notkestrasse 85, 22607 Hamburg, Germany. ^{*}These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.B. (email: Christian.Betzel@uni-hamburg.de) or C.W. (email: cwrenger@icb.usp.br)

Ideally, new antibacterial drugs should specifically target metabolic processes that are exclusive to bacteria (like the vitamin B1 biosynthetic pathway, which is conserved among most bacteria, plants and lower eukaryotes⁹) but absent in humans. Humans have to rely on dietary uptake of thiamin and thus side effects of potential new antibiotics targeting this vitamin metabolism are in principle avoided¹⁰. Thiamin pyrophosphate (TPP), the active form of thiamin, is essential for the metabolism of carbohydrates and the amino acid metabolism¹¹. In *Sa* six enzymes, ThiM, ThiD, ThiE, a nonspecific GTPase, TenA and TPK¹², are involved in the pathway (see Supplementary Fig. S1) forming TPP. Only the 3D structure of TenA (PDB code: 4FN6) is known so far¹³, consequently more detailed structural information about the *Sa* thiamin biosynthesizing enzymes is required to support structure based drug discovery investigations.

Limiting the availability of this essential co-factor will thus effectively inactivate a diverse spectrum of bacterial targets and will cause a self-multiplying downstream effect. Hence targeting this metabolism is similar in concept to developing a suicide drug¹⁴. This approach uses a special mechanism closely related to the pro-drug concept but with a unique extension: it uncouples the drug's sites of infiltration and action. As a result thiamin dependent pathways will be affected by the metabolized products and selective pressure will be moved from a single drug target towards several pathways depending on the activated form of thiamin.

Here we report a detailed structure-function analysis of the kinase ThiM, which phosphorylates THZ in one of the first steps in vitamin B1 biosynthetic pathway. In an attempt to achieve a maximum in accuracy and reliability of the structure, micro crystals grown under microgravity conditions onboard Chinese space mission Shenzhou-8¹⁵ were used to prepare seed stock solutions to grow crystals suitable for X-ray diffraction. In addition to the native structure, complexes with THZ and two selected lead compounds were analyzed.

The goal of these investigations was to identify potential pro-drug THZ analogues that do not inhibit ThiM, but will be accepted as substrates, yielding non-functional TPP analogues. In consequence these products cannot act as co-factors and the pathogen's metabolic homeostasis will be disrupted, leading to energy imbalance, growth stoppage and, ultimately, death. As the sulfur atom in the thiazolium ring of thiamin is indispensable for the stabilization of a reactive carbanion at C2 position, forming an ylid-group, two THZ analogues with pyrazole and imidazole structure and lacking sulfur in the heterocycle were selected for comparative complex structure analyses and, in support of these studies, *in vitro* analyses of biochemical function were performed for them as well. It had already been shown that thiamin derivatives having an oxazolium or imidazolium ring are less reactive than thiazolium analogues, since in absence of 3d orbitals the transition state cannot be stabilized^{16–18}. These comparative structural analyses of active site regions in combination with the data obtained from enzymatic assays yielded essential information to further design and optimize potential pro-drug THZ analogues.

Results

SaThiM structure. The quaternary structure of *Sa*ThiM is a triangular homo-trimer with dimensions of approximately 72 Å on a side and 47 Å thick with three independent active sites, each located within interface regions between two monomers (see Figs 1 and 2). The structure of native *Sa*ThiM was solved and refined to 2.1 Å resolution with R and R_{free} values of 20.34% and 23.04% respectively. Data collection and refinement parameters are summarized in Table 1 and the coordinates and data are available in PDB 5CM5. The homologue structure of *Bacillus subtilis* ThiK (PDB code: 1C3Q¹⁹), sharing 38% sequence identity, was used as a search model for molecular replacement phasing.

The secondary structure of *Sa*ThiM consists of approximately 50% α -helical and approximately 18% β -sheet structural elements (see Fig. 1). The core of each monomer contains nine β -sheets that are flanked by five α -helices and six α -helices. The first six β -sheets are parallel to each other with a corresponding strand order 2-1-3-4-5-6. The following β -sheet strand 7 is antiparallel to strand 6 and following β -sheet strand 8, which is antiparallely flanked by β -sheet strand 9. This fold of the phosphotransferase ThiM is homologous to that

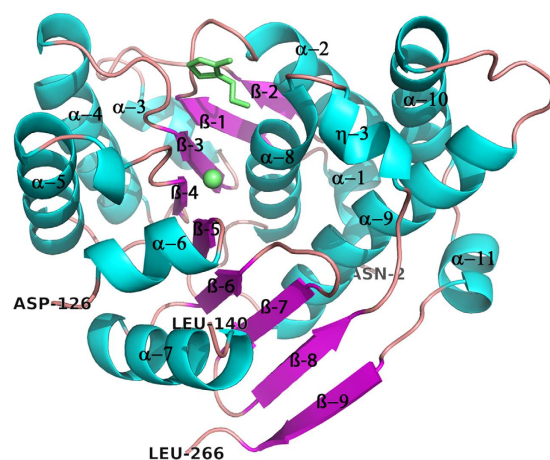


Figure 1. Overall structure of a ThiM monomer. Cartoon representation of the *Sa*ThiM monomer. Helices are coloured in cyan and β -sheets in magenta. The active site region with a bound THZ and a magnesium ion are highlighted and colored in green.

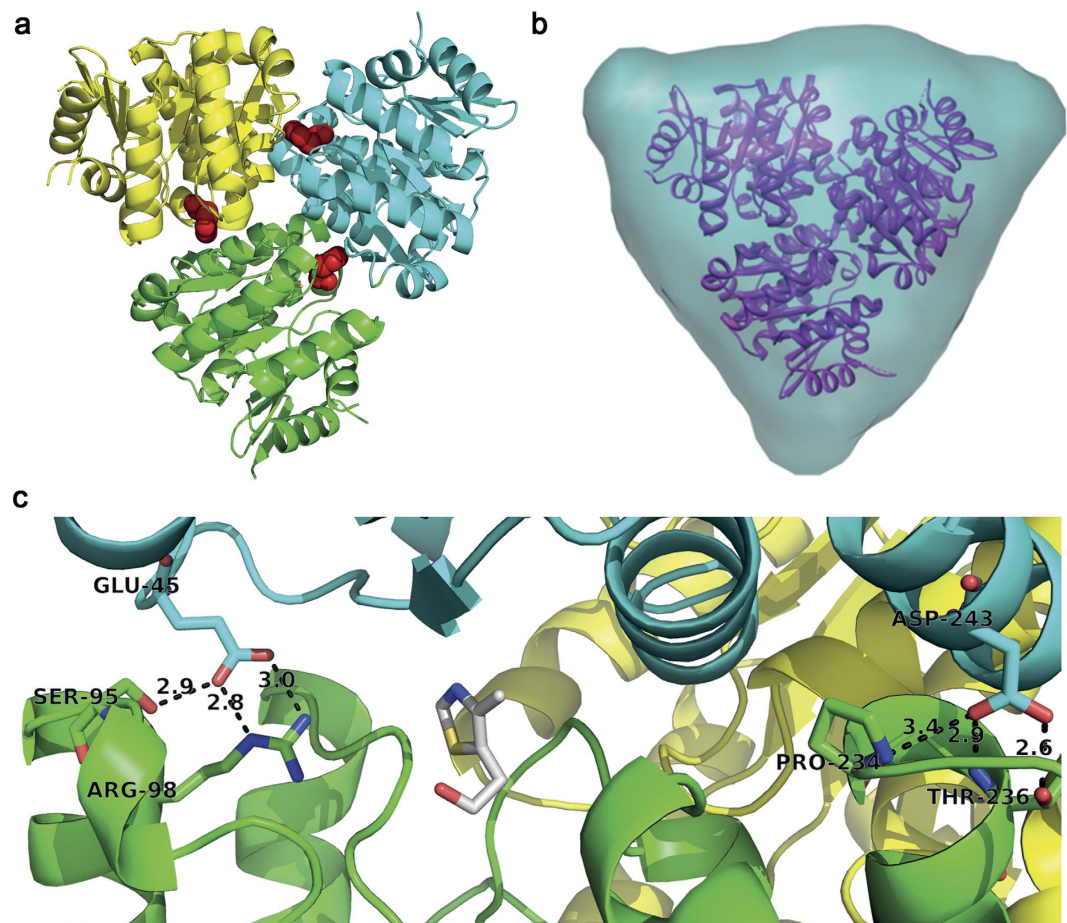


Figure 2. Structure of the ThiM trimer and its interface. (a) Cartoon representation of the *SaThiM* trimer, each monomer colored differently. The three active sites are located within the interface regions of the monomers, bound THZ molecules are shown as red spheres. (b) Front view of the *SaThiM* trimer *ab initio* shape in turquoise obtained from SAXS measurements and superimposed on the crystal structure, shown in a purple ribbon representation. (c) Detailed view to the interface and active site region with functional residues shown in stick representation with carbon chain atoms in the respective chain color, nitrogen in blue and oxygen in red. THZ is also shown in stick representation with carbon atoms in grey, nitrogen in blue and oxygen in red.

common to the group of ribokinase-like folded kinases²⁰. For all six molecules in the asymmetric unit no electron density was observed for the last two residues of the C-terminus and for the loop region between residues 126 and 140. Therefore these residues were not modeled. Analysing the interface regions of protomers A, B and C revealed that 10 intermolecular H-bonds support the complex formation. Analysis was done by manual inspection using the program PDBsum²¹. All intermolecular interactions are summarized in Supplementary Tables S1 and S2. The total surface of one monomer was calculated to be approximately 11,500 Å² with almost 1,000 Å² buried upon assembly of two monomers. The ThiM trimer has approximately 27,000 Å² solvent accessible surface area, and approximately 5,700 Å² of the monomeric surface accessible area is buried upon quaternary structure formation (see Fig. 2).

The trimeric assembly observed in the crystal structure was validated in solution with small angle X-ray scattering (SAXS). The experimental radius of gyration R_g of 30 ± 2 Å, the maximum dimension D_{max} of 90 ± 10 Å and the shape of the $p(r)$ function (Supplementary Fig. S2), point to a globular structure. The protein molecular mass estimated from the Porod volume and from the *ab initio* model built with the program DAMMIF²² are 91 ± 8 kDa and 77 ± 8 kDa respectively, which is compatible with trimeric *SaThiM* in solution. The DAMMIF model with enforced P3 symmetry is well superimposable with the crystal structure and resembles the trimeric structure in solution, as shown in Fig. 2b. The scattering computed from the trimeric model where the missing residues were symmetrically reconstructed with the program CORAL²³ fits the experimental SAXS data well with $\chi^2 = 0.6$ (see Supplementary Fig. S2). All SAXS-derived parameters are summarized in Table 2, the scattering data and the models are deposited in SASBDB²², code: SASDAX8.

A sequence alignment of ThiM with homologous structures by means of a NCBI/Blast PDB search option showed only a moderate degree of structural conservation among ThiM-type enzymes in bacteria (see Supplementary Fig. S3). *SaThiM* shares 32% amino acid sequence identity with *E. faecalis* ThiM (PDB code: 3DZV), 38% with *B. subtilis* ThiK (PDB code: 1C3Q) and 39% identity with *P. horikoshii* ThiM (PDB code:

	ThiM native	ThiM-THZ	ThiM-cpd 1	ThiM-cpd 2
Data collection				
Space group	P2 ₁	P1	P1	P1
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	62.6, 103.5, 126.2	62.6, 62.7, 108.5	62.0, 62.4, 108.3	62.4, 62.5, 109.2
α , β , γ (°)	90.0, 99.5, 90.0	92.2, 91.4, 101.3	92.6, 91.4, 101.3	92.6, 92.1, 101.5
Resolution (Å)	20.0–2.09	20.0–1.90	30.0–1.87	30.0–1.62
<i>R</i> _{merge}	7.8 (49.2)	2.7 (21.1)	6.9 (54.9)	6.9 (51.3)
<i>I</i> / σ	17.3 (4.2)	21.8 (4)	12.7 (2.5)	10.7 (2.5)
Completeness (%)	99.9 (100)	94.7 (92.9)	95.8 (94.7)	93.4 (90.5)
Redundancy	7.8 (7.7)	2 (1.9)	3.6 (3.5)	3.6 (3.6)
Refinement				
Resolution (Å)	20.0–2.09	20.0–1.90	30.0–1.87	30.0–1.62
No. reflections	88976	117078	119598	181567
<i>R</i> _{work} / <i>R</i> _{free}	20.34/23.04	18.61/20.95	16.86/19.90	18.24/20.28
No. atoms				
Protein	11474	11149	11228	11453
Ligand	–	54	66	54
Ions	–	4 Mg ²⁺	4 Mg ²⁺	5 Mg ²⁺
Water	552	371	486	511
<i>B</i> -factors				
Protein	39.1	41.1	37.5	30.5
Ligand	–	43.2	38.1	24.9
Ion	–	41.2	44.8	43.6
Water	33.4	26.0	33.8	26.1
R.m.s. deviations				
Bond lengths (Å)	0.016	0.017	0.019	0.017
Bond angles (°)	1.615	1.598	1.738	1.699

Table 1. Data collection and refinement statistics for native SaThiM, SaThiM in complex with THZ and in complex with two substrate analogues, cpd 1 and cpd 2. *Values in parentheses are for highest-resolution shell.

Data collection parameters	
Beamline	EMBL beamline X33 (DORIS III, DESY) ⁴⁸
Wavelength (Å)	1.5
<i>s</i> range (Å ⁻¹)	0.01–0.6
Exposure time (s)	120
Concentration range (mg/ml)	1.3–9.9
Temperature (°C)	12
Structural parameters	
<i>R</i> _g (Å) (from Guinier)	30 ± 2
<i>R</i> _g (Å) (from <i>p</i> (<i>r</i>))	29 ± 3
<i>D</i> _{max} (Å)	90 ± 10
Porod volume estimate (Å ³)	145,000 ± 15,000
Molecular mass from Porod volume (kDa)	91 ± 9
Molecular mass from dummy atom modelling (kDa)	77 ± 8
Monomeric molecular mass calculated from sequence (kDa)	29.8
Software employed	
Primary data reduction	SASFLOW pipeline ⁴⁹
<i>Ab initio</i> analysis	DAMMIF ⁴⁵
Rigid body modelling	CORAL ²³

Table 2. SAXS data collection and derived parameters.

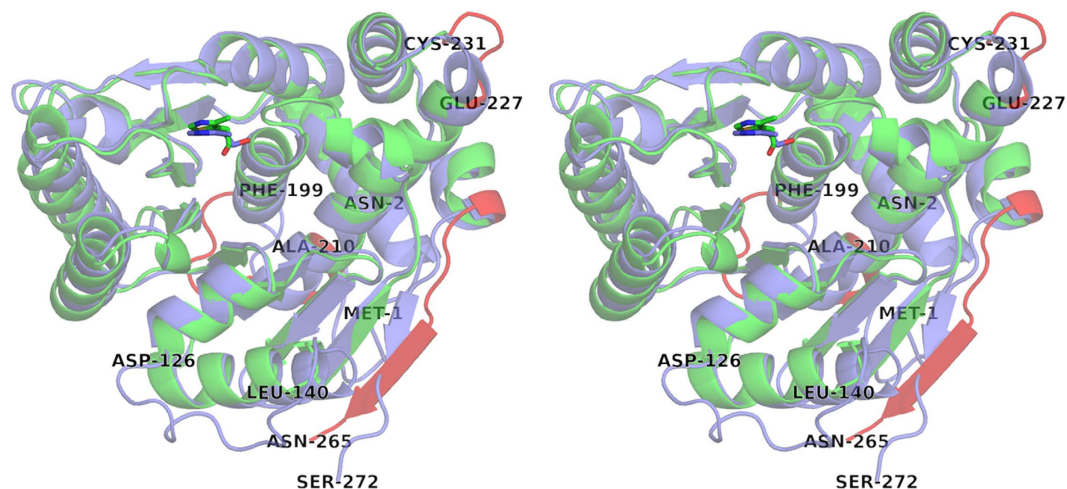


Figure 3. Superimposed structures of SaThiM and BsThiK. Stereo view of the superimposed homologues structures of SaThiM (colored in green) and ThiK from *Bacillus subtilis* (colored in blue) in a cartoon representation. The regions with higher differences, the loop region from residues 199–210 connecting helix 8 and 9 and close to N- and C-termini, are colored in red. The loop regions 126–140 are not modeled into SaThiM structure.

3HPD). The corresponding C α r.m.s.d. values are 1.6, 1.3, and 1.2 Å and maximum displacements were calculated to be 4.6, 4.5 and 4.2 Å, using the protein structure comparison service Fold at the European Bioinformatics Institute²⁴. The overall fold of the structures is conserved, whereas higher differences were found in the loop regions from residues 126–140 (connecting helix 6 and 7), 199–210 (connecting helix 8 and 9), 227–231 (connecting helix 9 and 10) and the regions close to N- and C-termini (see Figs 1 and 3). In Sa, part of the interface stabilization is conserved and mediated through inter-protomer H-bonds: E45 forming three H-bonds to R98, and S95 and D243 forming three H-bonds to P234 and T236 in the neighboring monomer (see Fig. 2c). A comparison of homologous hydroxyethylthiazole kinases from *B. subtilis* and *P. horikoshii* shows that this pattern of stabilization in the interface regions of the trimeric assembly, with two defined positions with amino acids forming two H-bonds to two neighboring amino acids, is highly conserved.

Active site and enzyme kinetics. Based on the crystal structure in complex with the natural substrate THZ, refined to 1.90 Å resolution with R and R_{free} values of 18.61% and 20.95% respectively (Table 1, PDB: 5COJ), the three identical active sites of the trimer were characterized and the mode of action for phosphorylation could be proposed. In the crystal structure, THZ identically occupies all six active sites of the two trimers in the asymmetric unit. A view of the experimental electron density of ThiM in complex with THZ can be found as Supplementary Fig. S4. THZ binding is mediated by interactions with amino acids N19, V21, G61, V90, T186, G187 and C190 of one monomer and P37, A38 and M39 of the corresponding neighboring monomer, all located in the interface region. Substrate binding is mainly stabilized by H-bonds formed by the nitrogen of the heterocycle and via the flexible hydroxyl group of THZ (see Fig. 4a). The orientation of the hydroxyl group, as extracted from the THZ complex formation in all six active sites in the asymmetric unit, can adopt slightly different orientations (see Supplementary Table S3). In summary THZ forms six direct H-bonds to residues M39, P37, G61, T186 and C190. For the ThiM-THZ complex four Mg²⁺ ions, essential for the phosphor transfer reaction, were located within four active sites of the two trimers in the asymmetric unit. The Mg²⁺ ions are coordinated by residues D88, K115 and E120 and one solvent water molecule as shown in Fig. 4a. No Mg²⁺ ion was found in the native structure. Utilizing also the structural data obtained from homologous ThiM-complex structures, such as the mutated ThiK homologue, from *Bacillus subtilis*¹⁹ (PDB code: 1ESQ) with bound ATP and THZ, as well data known about conserved phosphoryl transfer within ribokinases-like kinases²⁵, the following mode of action can be proposed, which is additionally illustrated in Supplementary Fig. S5. Enhanced nucleophilicity of the THZ alcohol is expected due to a proton-relay system consisting of a water molecule, which is coordinated by a magnesium ion and C190. Additionally two oxygen atoms of the ATP- γ -phosphate are coordinated by the magnesium ion and N117 resulting in an enhanced electrophilicity of the γ -phosphate and a pentavalent intermediate between the γ -phosphate and THZ. Further, the transition state is stabilized by hydrogen bonds of T160 OG to the α -phosphate and of K115 NZ to the β -phosphate.

In order to identify substrate analogues for SaThiM, a structure-based virtual screening approach was conducted. At first, a focused compound library was generated by searching the ZINC database²⁶ for THZ structure analogues by means of a SMARTS query containing several restrictions: i) a 5-membered heterocyclic ring including a nitrogen atom to enable hydrogen bonding with N39 and excluding a sulfur atom in order to inhibit the carbanion formation for thiamin-dependent enzymes²⁷, and ii) a hydroxy-ethyl moiety in position 4 of the 5-membered ring to allow phosphorylation by the enzyme. Applying these constraints resulted in 78 compounds that were docked into the SaThiM active site using AutoDock 4.2.3²⁸ with default settings. Based on docking scores and visual analysis of compound-binding modes, 9 compounds were identified as putative SaThiM

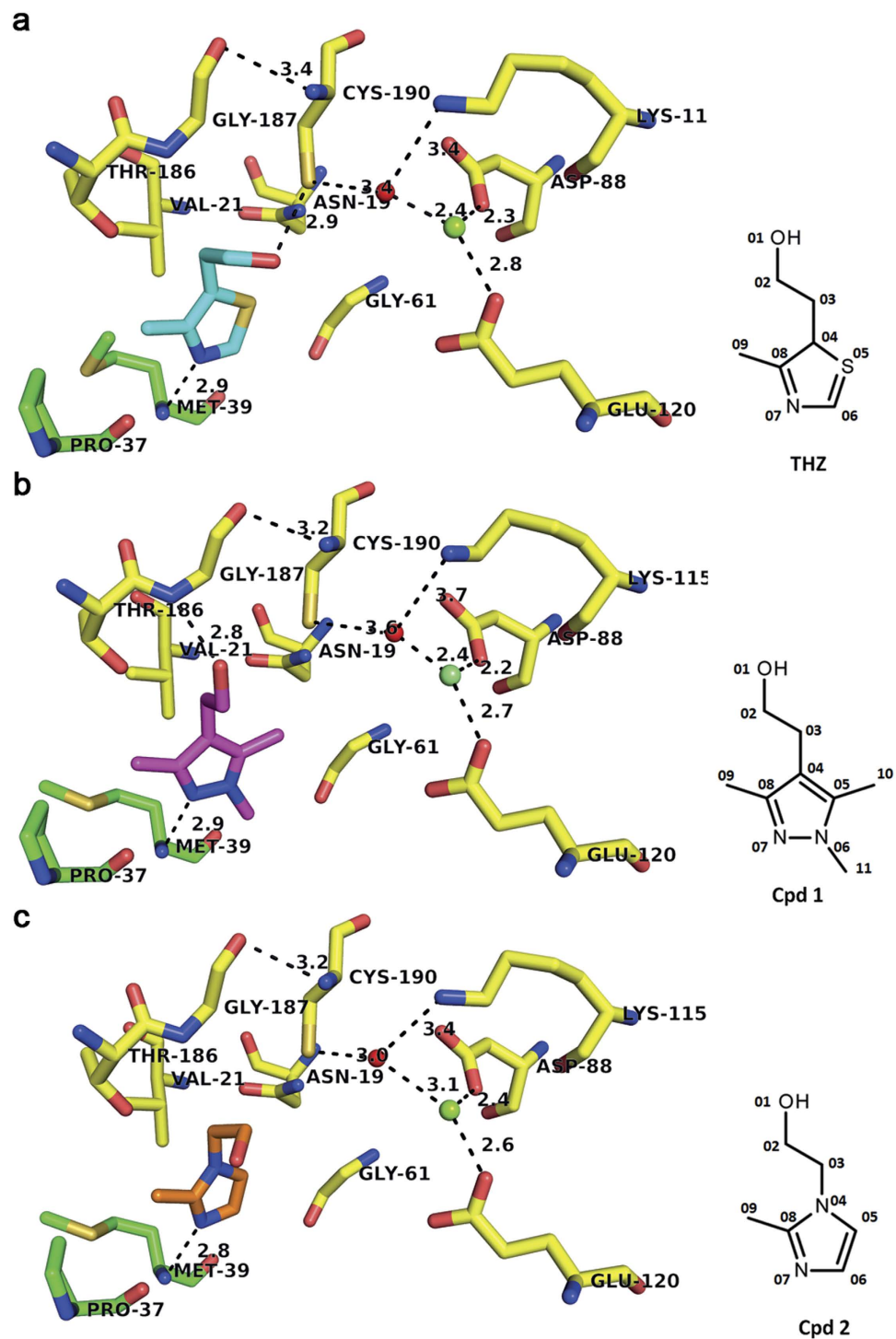


Figure 4. *SaThiM* active site with bound substrates. *SaThiM* active site with bound THZ (a), cpd 1 (b) and cpd 2 (c). The active site residues are shown in stick representation with carbon chain atoms in the respective chain color, nitrogen in blue, oxygen in red and sulphur in yellow. For THZ carbon atoms are colored in cyan, for cpd 1 in magenta and for cpd 2 in orange. Waters are displayed as red and Mg^{2+} ions as green spheres. H-bonds are shown as dashed lines. Consensus nomenclature of the substrate THZ and selected lead compounds cpd 1 and cpd 2 are shown, respectively.

substrates. Out of this *in silico* selected set of molecules, two were chosen as the most promising analogues for further investigations: 2-(1,3,5-Trimethyl-1H-pyrazol-4-yl) ethanol, hereinafter referred to as compound 1 (cpd 1) and 2-(2-Methyl-1H-imidazol-1-yl) ethanol referred to as compound 2 (cpd 2). The activity of ThiM was analyzed *in vitro*, as described in the methods section, showing that the specific activities of cpd 1 and cpd 2 are

	Specific Activity [nmol/min/mg]	K_M [μ M]	k_{cat} [min^{-1}]
THZ ¹²	4880 \pm 488	44 \pm 5	137 \pm 13
cpd 1	7297 \pm 267	834 \pm 147	215 \pm 8
cpd 2	7418 \pm 91	831 \pm 169	218 \pm 3

Table 3. Kinetic parameters of *Sa*ThiM phosphorylating THZ¹², cpd 1 and cpd 2.

approximately 1.5 times higher compared to THZ, however show only 1/10 of the efficiency (k_{cat}/K_M) in comparison to the natural substrate THZ (see Table 3 and Supplementary Fig. S6). This is also reflected by the higher K_M -value of the compounds.

In summary the kinetic data confirmed that the selected compounds were accepted as substrates, even if *in vitro* turnover was slightly lower, and could thus be used for further investigations and to optimize effective pro-drug compounds targeting the vitamin B1 biosynthetic pathway. Based on these findings complexes of *Sa*ThiM with cpd 1 and cpd 2 were prepared and crystallized.

The structures of these complexes could be refined to 1.87 Å and 1.62 Å resolution with R and R_{free} values of 16.86% and 19.90% for the cpd 1 complex and R and R_{free} values of 18.24% and 20.28% for the cpd 2 complex. Data collection and refinement parameters are summarized in Table 1 and the coordinates for the structures, as well as the experimental diffraction amplitudes have been deposited at the Protein Databank (<http://www.rcsb.org>) with entry codes 5CGA and 5CGE. The chemical structure as well as the compounds bound in the active site are shown in Fig. 4b,c, respectively. The overall coordination of the THZ analogue compounds in the active site regions of the ThiM trimer is remarkably comparable to that of the of natural substrate THZ (see Fig. 4a). Both compounds form a H-bond to the amide nitrogen of amino acid M39 and show a similar orientation of their hydroxyethyl groups. The thiazole moiety was replaced by the non-sulfur containing heterocycles pyrazole and imidazole for cpd 1 and cpd 2, respectively. Moreover, cpd 1 contains two additional methyl groups at position 5 and 6 of the heterocycle (see Fig. 4b,c, respectively). These structural differences to the natural substrate result in a 1.5 higher specific activity for both cpd 1 and cpd 2, even if the location of the compounds in the active site is highly comparable to that of the natural substrate, as shown in Table 3, confirming the appropriate pre-positioning for the phosphoryl transfer reaction.

Discussion

The shortfall in the development of new antibiotics and the misguided overuse of the currently available antibiotics has caused an unfortunate spread of multi-resistant bacteria in recent years. Because of this, rational structure-based drug discovery is urgently required⁶ to cover the loss by supporting accelerated design of highly selective antibiotics with an elaborated mode of action. To our knowledge this study represents the first approach in targeting the *Sa* thiazole kinase ThiM of the bacterial vitamin B1 pathway, applying structure based pro-drug discovery. In order to improve crystal quality of *Sa*ThiM, we performed crystallization trials under microgravity conditions in combination with micro seeding techniques. This procedure has previously been applied successfully to optimize crystal quality²⁹. We observed an increase in crystal quality compared to native triclinic *Sa*ThiM crystals obtained in previous experiments³⁰. Soaking of the respective crystals with ThiM substrate and selected pro-drug compounds resulted in loss of crystal symmetry. Similar observations were made for crystals of the ThiM homologue ThiK from *B. subtilis*, where soaking with substrates led to crystalline complexes with monoclinic symmetry, in comparison to native rhombohedral crystals¹⁹. This effect could be attributed to a slightly changed fold in the loop region 126–140, which changed the interactions and crystal packing of the trimeric subunits in the asymmetric unit. *Sa*ThiM revealed a conserved ribokinase-like fold and a trimeric assembly, which was confirmed by complementary SAXS experiments. The three active sites, each located within the interface regions of two monomers respectively, are homologous to other known phosphoryl transfer ribokinase-like kinases²⁵. Based on the overall structural information of ThiM complexes obtained, we propose a catalytic pathway of *Sa*ThiM. Further, we identified the role of C190 and the coordinated magnesium ion in enhancing the nucleophilicity of the THZ alcohol and thereby facilitating its phosphorylation via a proton relay mechanism. Structural data of THZ bound in the active sites of *Sa*ThiM allowed *in silico* screening and identification of two initial lead compounds. In terms of the pro-drug development concept, these compounds were used to prepare enzyme-compound complexes. The results of the biochemical assays, complemented by the structural data, support the suggested pathway of phosphorylation. Comparing the specific activities of ThiM in presence of the two compounds revealed a similar turnover as that of the natural substrate THZ, which clearly validates the approach to combine complementary methods for pro-drug discovery in order to identify and improve lead compounds.

The two novel compounds applied for co-crystallization in this study lack the thiazolium sulfur, present in the natural thiazolium moiety of thiamin, which is replaced by a carbon atom, supporting slightly higher flexibility in the active site in consequence of the reduced van der Waals radius. Overall, the orientation of both compounds in the active sites of ThiM is very similar to that observed for the natural substrate THZ.

For both compounds the N3 is retained and a second nitrogen atom at position 2 and 5 is included in the heterocycle of cpd 1 and cpd 2 respectively. Moreover, cpd 1 has two additional methyl groups bound to atoms 1 and 2, which can provide more hydrophobic interactions in binding co-factor dependent enzymes after complexation to the HMP moiety by ThiE thereby enabling the cpd 1-HMP moiety to act as a competitive inhibitor. Cpd 2 has an additional nitrogen atom in the heterocycle and could, like cpd 1, undergo a TPP imino tautomerization after fusion with HMP. As a result this substance could be channeled via thiaminase (TenA) into the bacterial thiamin pool, where it could act as a thiamin antagonist equally potent as pyrithiamine³¹. The data presented

here summarize the first steps towards the development of thiamin analogue compounds, only produced *in vivo* in bacteria cells, which can inhibit preferably multiple co-factor dependent enzymes and in addition might also block the thiamin riboswitch in the bacteria³². This approach will open a route for the development of new antimicrobial substances to treat MRSA infections.

Methods

Purification and Crystallization. The expression, purification and initial crystallization of native SaThiM has been described previously³⁰. To optimize the crystal quality and try for a better crystal packing, crystallization experiments under microgravity conditions were performed in the DLR SIMBOX experiment^{15,33} onboard the Chinese space mission Shenzhou-8¹⁵. The SIMBOX counter diffusion crystallization experiments were coordinated by the Institute of Biophysics of the Chinese Academy of Sciences, Beijing. For the space experiment, protein solutions were dialyzed overnight against 100 mM Tris, 150 mM NaCl, pH 8.0. ThiM was concentrated to approximately 20 mg/mL, an equal volume of 0.5 M magnesium formate was added at 4 °C, incubated for at least 6 h and centrifuged at $20,000 \times g$ for 30 min. Prior to the microgravity experiments, crystallization conditions were extensively optimized under laboratory conditions, including transport simulations. The final protein solutions were analyzed by dynamic light scattering and showed a stable monodisperse peak for several days. The capillary length of the SIMBOX crystallization hardware was 10 mm with an inner diameter of 1 mm. The capillary was filled with a ThiM solution using a concentration of 12 mg/mL and 2% (w/v) low melting point agarose (Serva) was used to seal the capillaries. Subsequently, the capillaries were inserted carefully into a plastic casing, containing four separated sections filled with foam material and sealed with a metal lid. The foam was saturated with approximately 300 μ L of precipitant solution (20% PEG 3,350 (w/v), 7% isopropanol (v/v)). For the microgravity experiment 4 capillaries were used and positioned in the DLR SIMBOX. The microgravity experiment was performed at a stable temperature of 20 °C for 17 days (30.10.2011–17.11.2011) onboard the Shenzhou-8 space mission. After return of the experiment the capillaries were inspected directly after arrival in Beijing. The capillaries were opened at one end and the agarose was removed carefully. Obtained micro-crystals were used to prepare native ThiM and complex crystals applying hanging drop vapor diffusion and a streak seeding procedure³⁴. Seed stocks were prepared following the manufacturer's instructions (Jena BioScience Beads-for-seeds). 24-well CPL-130 plates (Jena Bioscience) for hanging drop vapor diffusion experiments were filled with 500 μ L of precipitant solution and ThiM was used at a concentration of 10 mg/mL. In 48-well MRC sitting drop plates (Molecular Dimensions) for sitting drop vapor diffusion, two rows of the precipitants with different PEG 3,350 concentrations were prepared by filling 50 μ L of precipitant (18, 20 and 22% PEG 3,350 (w/v), 0.2 M magnesium formate and 5% isopropanol (v/v)) into the reservoirs. Purified ThiM was dialyzed overnight against 50 mM Tris (pH 8.0), 150 mM NaCl and directly used at 10 mg/mL concentration after centrifugation ($100,000 \times g$, 4 °C for 50 min). To obtain a ThiM-THZ complex native crystals were supplemented with THZ at a concentration of 0.5 mM overnight. To obtain ThiM-cpd 1 and ThiM-cpd 2 complexes, crystals were soaked with a 25 mM and 20 mM compound 1 and 2 solutions respectively overnight. Prior to data collection crystals were soaked in precipitant solution containing 10% v/v glycerol as cryoprotectant and THZ, cpd 1 or cpd 2 respectively.

Structure analysis. Diffraction data of native ThiM were collected as described before³⁰. A complex crystal of ThiM-THZ was used to collect diffraction data using a MARCCD165 detector at the DORIS consortium beamline X13 (HASYLAB/DESY) using a wavelength of 0.81 Å at 100 K. Diffraction data of ThiM-cpd 1 and ThiM-cpd 2 crystals were collected at the EMBL beamline P14 (PETRAIII, DESY, Hamburg) at 0.98 Å wavelength and 100 K with a PILATUS 6M detector.

Data processing and scaling of native ThiM and ThiM-THZ were carried out using DENZO and SCALEPACK³⁵, ThiM-cpd 1 and ThiM-cpd 2 data sets were processed applying the program XDS³⁶. The uncomplexed crystals belonged to the space group $P2_1$ (Unit cell parameters: $a = 62.6$ Å, $b = 103.5$ Å, $c = 126.2$ Å and $\beta = 99.5^\circ$) with six ThiM molecules, forming 2 trimers, in the asymmetric unit. The packing parameter V_M was calculated to be $2.3 \text{ \AA}^3/\text{Da}^{37}$, which corresponded to a solvent content of 47%. Native ThiM crystals, which were grown applying microgravity grown seed stock crystals, were in the monoclinic space group $P2_1$. Soaked THZ, cpd 1 and cpd 2 complex crystals are assigned to the triclinic space group $P1$, due to slightly changed crystal contacts of the six ThiM molecules in the unit cell. Unit cell parameters and data collection and scaling statistics are summarized in Table 1. The coordinates of *Bacillus subtilis* ThiK (PDB code: 1C3Q) were used as a starting model for molecular replacement applying the program PHASER³⁸ within the CCP4 suite³⁹. The refined native ThiM structure was subsequently used as a starting model for the complex structures, ThiM-THZ, ThiM-cpd 1 and ThiM-cpd 2 applying MOLREP⁴⁰. The models were refined through iterative cycles of restrained refinement by REFMAC5⁴¹ and phenix.refine⁴² and manual rebuilding applying the program COOT⁴³.

Resolution limits of the diffraction data were set applying the criteria $I/\sigma > 2$, monitoring R_{merge} with a limit of approximately 50%. These criteria resulted in resolution limits of 2.1 Å of the native diffraction data set, 1.9 Å for the THZ complex, 1.85 Å for the cpd 1 and 1.6 Å for the cpd 2 complex. After first manual model building and refinement solvent molecules were included applying Fo-Fc difference maps. For the complex crystals the THZ, cpd 1 and cpd 2 moieties were introduced and refined. Optimized geometry restraints for bond distance and angles of THZ as well as for cpd 1 and cpd 2 were generated with Phenix eLBOW⁴² and applied for refinement. Magnesium ions, if present, were positioned applying Fo-Fc density maps contoured at 2.5σ in a tetragonal coordination with ligand residues L115, E120 and D88.

For native ThiM 98% of residues are in the favoured, 2% are in the additional allowed regions of the Ramachandran plot and no residue was found in the disallowed regions. For ThiM-THZ, ThiM-cpd 1 and ThiM-cpd 2 99% are in the favoured regions and 1% is in the additional allowed regions of the Ramachandran plot and no residue was found in the disallowed regions. Analysis was done with MolProbity⁴⁴.

Small angle X-ray scattering. SAXS data of native ThiM were collected at the EMBL beamline X33 (DORIS III, DESY, Hamburg)⁴⁵. Using a 1M-W PILATUS detector at a sample-to-detector distance of 2.7 m and a wavelength of $\lambda = 1.5 \text{ \AA}$, the range of momentum transfer $0.01 < s < 0.6 \text{ \AA}^{-1}$ was covered ($s = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle). Four solute concentrations ranging between 1.3 and 9.9 mg/ml were measured at 12 °C. To monitor for the radiation damage, eight successive 15-second exposures of protein solutions were compared and no statistically significant changes were detected. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the buffer was subtracted and the difference curves were scaled for protein concentration. The low angle data measured at lower protein concentrations were merged with the higher concentration data to yield the final composite scattering curves. The radius of gyration R_g , the Porod volume of the hydrated particle along with the distance distribution function $p(r)$ and the maximum intra-molecular distance D_{max} were derived using the automated SAXS data analysis pipeline SASFLOW⁴⁶. The molecular mass (MM) was obtained from the Porod volume: for globular proteins Porod volumes in \AA^3 are about 1.6 times the MMs in Da. *Ab initio* shape model was reconstructed using the bead modelling program DAMMIF⁴⁵. This program represents the particle shape by an assembly of densely packed beads and employs simulated annealing to construct a compact interconnected model fitting the experimental data. A molecular mass estimate is derived from the volume of the model. Ten DAMMIF runs were performed to check the stability of solution, and the results were well superimposable with each other. 15 C-terminal residues missing from the high-resolution X-ray structure were reconstructed with the rigid body modelling program CORAL²³, the fit to the experimental data was refined with CRY SOL⁴⁶. All SAXS-derived parameters are summarized in Table 2, the scattering data and the models are deposited in SASBDB²², code: SASDAX8.

Activity analysis. *Sa*ThiM activity assays were performed in 100 μl buffer solution containing 50 mM potassium phosphate (pH 7.5), 10 mM MgCl_2 , 4 mM ATP (supplied by Sigma), 50 nCi ATP- $[\text{P}^{33}]$ and 0–4 mM THZ. 315 ng of purified protein was added and incubated for 15 min at 37 °C. Reactions were quenched at 95 °C for 2 minutes and reaction products were analyzed as described before¹². *De novo* synthesized THZ- $[\text{P}^{33}]$ was quantified using a LS5000 CE scintillation detector (BeckmanCoulter). The kinetic parameters were calculated using the GraphPad Prism 5 software (GraphPad Software) as described previously⁴⁷ and summarized including the standard deviation in Table 3 and as Supplementary Fig. S6.

References

- Lowy, F. D. Staphylococcus aureus infections. *N. Engl. J. Med.* **339**, 520–32 (1998).
- Lowy, F. D. Antimicrobial resistance: The example of Staphylococcus aureus. *J. Clin. Invest.* **111**, 1265–1273 (2003).
- Otto, M. Basis of virulence in community-associated methicillin-resistant Staphylococcus aureus. *Annu. Rev. Microbiol.* **64**, 143–62 (2010).
- Spellberg, B., Powers, J. H., Brass, E. P., Miller, L. G. & Edwards, J. E. Trends in antimicrobial drug development: implications for the future. *Clin. Infect. Dis.* **38**, 1279–1286 (2004).
- Wenzel, R. P. The antibiotic pipeline—challenges, costs, and values. *N. Engl. J. Med.* **351**, 523–6 (2004).
- Silver, L. L. Challenges of antibacterial discovery. *Clin. Microbiol. Rev.* **24**, 71–109 (2011).
- Chambers, H. F. & DeLeo, F. R. Waves of resistance: Staphylococcus aureus in the antibiotic era. *Nat. Rev. Microbiol.* **7**, 629–641 (2009).
- Secretariat of the WHO. Draft global action plan on antimicrobial resistance - Report by the Secretariat. Sixty-seventh world health assembly, **A67/39**, Add.1, 1–3 (08 May 2014). http://apps.who.int/gb/ebwha/pdf_files/WHA67/A67_39Add1-en.pdf.
- Jurgenson, C. T., Begley, T. P. & Ealick, S. E. The structural and biochemical foundations of thiamin biosynthesis. *Annu. Rev. Biochem.* **78**, 569–603 (2009).
- Begum, A., Drebes, J., Perbandt, M., Wrenger, C. & Betzel, C. Purification, crystallization and preliminary X-ray diffraction analysis of the thiaminase type II from Staphylococcus aureus. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **67**, 51–3 (2011).
- Pohl, M., Sprenger, G. a & Müller, M. A new perspective on thiamine catalysis. *Curr. Opin. Biotechnol.* **15**, 335–42 (2004).
- Müller, I. B. *et al.* The vitamin B1 metabolism of Staphylococcus aureus is controlled at enzymatic and transcriptional levels. *PLoS One* **4**, e7656 (2009).
- Begum, A. *et al.* Staphylococcus aureus thiaminase II: oligomerization warrants proteolytic protection against serine proteases. *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 2320–9 (2013).
- Drebes, J., Künz, M., Pereira, C. A., Betzel, C. & Wrenger, C. MRSA infections: from classical treatment to suicide drugs. *Curr. Med. Chem.* **21**, 1809–19 (2014).
- Preu, P. & Braun, M. German SIMBOX on Chinese mission Shenzhou-8: Europe's first bilateral cooperation utilizing China's Shenzhou programme. *Acta Astronaut.* **94**, 584–591 (2014).
- Haake, P. & Miller, W. B. A Comparison of Thiazoles and Oxazoles. *J. Am. Chem. Soc.* **85**, 4044–4045 (1963).
- Hafferl, W., Lundin, R. & Ingraham, L. L. Activated Hydrogens in Compounds Related to Thiamine. *Biochemistry* **2**, 1298–1305 (1963).
- Breslow, R. On the Mechanism of Thiamine Action. IV. I Evidence from Studies on Model Systems. *J. Am. Chem. Soc.* **80**, 3719–3726 (1958).
- Campobasso, N., Mathews, I. I., Begley, T. P. & Ealick, S. E. Crystal Structure of 4-Methyl-5- β -hydroxyethylthiazole Kinase from Bacillus subtilis at 1.5 \AA Resolution. *Biochemistry* **39**, 7868–7877 (2000).
- Cheek, S., Zhang, H. & Grishin, N. V. Sequence and structure classification of kinases. *J. Mol. Biol.* **320**, 855–881 (2002).
- Laskowski, R. A. *et al.* PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* **22**, 488–90 (1997).
- Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* **43**, D357–D363 (2015).
- Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
- Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2256–68 (2004).
- Kenyon, C. P., Roth, R. L., van der Westhuyzen, C. W. & Parkinson, C. J. Conserved phosphoryl transfer mechanisms within kinase families and the role of the C8 proton of ATP in the activation of phosphoryl transfer. *BMC Research Notes* **5**, 131 (2012).
- Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **52**, 1757–68 (2012).

27. Settembre, E., Begley, T. P. & Ealick, S. E. Structural biology of enzymes of the thiamin biosynthesis pathway. *Curr. Opin. Struct. Biol.* **13**, 739–747 (2003).
28. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–91 (2009).
29. McPherson, A. Effects of a microgravity environment on the crystallization of biological macromolecules. *Microgravity Sci. Technol.* **6**, 101–9 (1993).
30. Drebes, J., Perbandt, M., Wrenger, C. & Betzel, C. Purification, crystallization and preliminary X-ray diffraction analysis of ThiM from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **67**, 479–81 (2011).
31. Alston, T. A. & Abeles, R. H. Enzymatic conversion of the antibiotic metronidazole to an analog of thiamine. *Arch. Biochem. Biophys.* **257**, 357–62 (1987).
32. Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G. M. & Breaker, R. R. Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem. Biol.* **12**, 1325–1335 (2005).
33. Ruyters, G. & Braun, M. SIMBOX-Experimente an Bord des Raumschiffs Shenzhou-8 – Deutsch-chinesische Kooperation in biomedizinischer Weltraumforschung. *Flug u. Reisemed.* **20**, 200–204 (2013).
34. D'Arcy, A., Sweeney, A. Mac & Haber, A. Modified microbatch and seeding in protein crystallization experiments. *J. Synchrotron Radiat.* **11**, 24–26 (2004).
35. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307–326 (1997).
36. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–32 (2010).
37. Kantardjiev, K. A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12**, 1865–1871 (2003).
38. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
39. Project, C. C. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760–3 (1994).
40. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
41. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–67 (2011).
42. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–21 (2010).
43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–32 (2004).
44. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21 (2010).
45. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
46. Svergun, D., Barberato, C. & Koch, M. H. J. CRY SOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
47. Wrenger, C. *et al.* Vitamin B1 *de novo* synthesis in the human malaria parasite *Plasmodium falciparum* depends on external provision of 4-amino-5-hydroxymethyl-2-methylpyrimidine. *Biol. Chem.* **387**, 41–51 (2006).
48. Blanchet, C. E. *et al.* Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443 (2015).
49. Franke, D., Kikhney, A. G. & Svergun, D. I. Automated acquisition and analysis of small angle X-ray scattering data. *Nucl. Instruments Methods Phys. Res. Sect. A. Accel. Spectrometers, Detect. Assoc. Equip.* **689**, 52–59 (2012).

Acknowledgements

This work was supported by the DLR via grant 50WB1017, by the DFG Excellence Cluster CUI (Centre for Ultra Fast Imaging) and via grant WR124/2, BMBF (01DN13037), the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2013/10288-1), by the Hamburg Ministry of Science and Research and Joachim Herz Stiftung as part of the Hamburg Initiative for Excellence in Research, the Hamburg School for Structure and Dynamics in Infection (SDI). Additionally this work was supported by the University of Hamburg (UHH)/University of Sao Paulo (USP) co-funding scheme CG 83.353. Further we would like to thank Benjamin Dose, Robin Schubert, Yong Yu and Jun Liu.

Author Contributions

J.D. conducted biochemical experiments, crystallization and diffraction data collection of native ThiM and ThiM-THZ. M.K. optimized ThiM purification, conducted crystallization, diffraction data collection and structure analysis of cpd 1 and cpd 2 complex structures. R.J.E. and D.O. prepared microgravity crystallization experiments. Kinetic assays were performed by J.D., I.B.M. and B.W. performed *in silico* work. J.D. and A.K. conducted the SAXS experiments, A.K. and D.S. analyzed the SAXS data. H.C. facilitated the Shenzhou microgravity experiments. J.D., M.K., M.P., C.W. and C.B. designed the overall study and wrote the manuscript.

Additional Information

Accession codes: PDB: 5CM5 (native), 5COJ (THZ complex), 5CGA (cpd 1 complex) and 5CGE (cpd 2 complex); SASBDB₂₃ code: SASDAX8.

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Drebes, J. *et al.* Structure of ThiM from Vitamin B1 biosynthetic pathway of *Staphylococcus aureus* – Insights into a novel pro-drug approach addressing MRSA-infections. *Sci. Rep.* **6**, 22871; doi: 10.1038/srep22871 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>