SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
MOLECULAR CONFORMATION
X-RAY CRYSTALLOGRAPHY

Received
22 May 2014

Accepted
8 July 2014

Published
23 September 2014

Correspondence and requests for materials should be addressed to Z.R. (raozh@nankai.edu.cn)

* These authors contributed equally to this work.

Homotypic dimerization of a maltose kinase for molecular scaffolding

Jun Li[1]*, Xiaotao Guan[1]*, Neil Shaw[1,3], Weimin Chen[3], Yu Dong[1], Xiaoling Xu[4], Xuemei Li[1] & Zihe Rao[1,2,3]

[1]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing, 100101, China, [2]Structure Biology Laboratory, Tsinghua University, Beijing, 100084, China, [3]Tianjin Key Laboratory of Protein Science, College of Life Sciences, Nankai University, Tianjin, 300071, China, [4]Institute of Aging Research, School of Medicine, Hangzhou Normal University, 310036, China.

***Mycobacterium tuberculosis* (*Mtb*) uses maltose-1-phosphate to synthesize α-glucans that make up the major component of its outer capsular layer. Maltose kinase (MaK) catalyzes phosphorylation of maltose. The molecular basis for this phosphorylation is currently not understood. Here, we describe the first crystal structure of *Mtb*MaK refined to 2.4 Å resolution. The bi-modular architecture of *Mtb*MaK reveals a remarkably unique N-lobe. An extended sheet protrudes into ligand binding pocket of an adjacent monomer and contributes residues critical for kinase activity. Structure of the complex of *Mtb*MaK bound with maltose reveals that maltose binds in a shallow cavity of the C-lobe. Structural constraints permit phosphorylation of α-maltose only. Surprisingly, instead of a Gly-rich loop, *Mtb*MaK employs 'EQS' loop to tether ATP. Notably, this loop is conserved across all MaK homologues. Structures of *Mtb*MaK presented here unveil features that are markedly different from other kinases and support the scaffolding role proposed for this kinase.**

Trehalose is a non-reducing disaccharide found in bacteria, archaea, fungi, plants and some invertebrates[1,2]. Surprisingly, mammalian cellular systems show a complete lack of trehalose[3]. In mycobacteria, trehalose serves as a component of cell wall glycolipids, and plays an important role in cellular response to stress[3]. Notably, trehalose is constitutively present in the cytoplasm of mycobacteria and is constantly subject to turnover; exemplifying the importance of trehalose in mycobacterial metabolism[4]. A key conversion of trehalose to α-glucans for the synthesis of the outer capsular layer and possibly methylglucose lipopolysaccharides contributes to the pathogenecity of *Mycobacterium tuberculosis* (*Mtb*)[5,6]. At least three different pathways are known to exist for the synthesis of α-glucans. These include the GlgC-GlgA pathway[6], the α(1->4) glucosyltransferase (Rv3032) pathway[6] and the TreS-MaK-GlgE-GlgB pathway[7]. These pathways have been shown to compensate each other for malfunctions in the pathways that stall α-glucan synthesis. For example, Δ*GlgE*-Δ*TreS* mutants of mycobacterium with a deficient TreS-MaK-GlgE-GlgB pathway are viable[8]. When an additional loss-of-function mutation of Rv3032 was introduced in these mutants, they lost viability, highlighting the redundancy in pathways for synthesis of α-glucans[8].

Maltose kinase functions in the TreS-MaK-GlgE-GlgB pathway that has been shown to synthesize α-glucans[8]. In this pathway, trehalose is first isomerized to maltose by trehalose synthase (TreS). This conversion is reversible[9,10]. Next, maltose is phosphorylated to maltose-1-phosphate by maltose kinase (MaK) by expending a molecule of ATP. The phospho-activated disaccharide is a substrate for glycosyltransferase E (GlgE). GlgE uses maltose-1-phosphate to elongate α(1->4) linked glucan chains. The last enzyme of this pathway, glycosyltransferase B (GlgB), mediates α(1->6)-branching of the glucan chain[11] (Supplementary Fig. S1). Interestingly, inactivation of GlgE leads to accumulation of maltose-1-phosphate, the product of MaK, during growth in presence of trehalose[8]. The accumulation of maltose-1-phosphate has been shown to be the cause of rapid death of *Mtb in vitro* and in mice[8].

In vitro viability studies on *Mtb* subjected to transposon mutagenesis have helped identify *mak* as an essential gene for the growth of *Mtb*[12,13]. The product of MaK-mediated catalysis, maltose-1-phosphate, is used for synthesis of α-glucans that make up almost 80% of the components of the outer capsule of *Mtb*[14]. However, unlike *Mtb*, micro-organisms like *Actinoplanes missouriensis* and *Streptomyces coelicolor* that express MaK homologues do not produce outer capsular layer. In addition, MaK is constitutively expressed in *Actinoplanes missouriensis*[15]. These observations suggest the possibility of additional roles for maltose-1-phosphate other than synthesis of capsular α-glucans. Biophysical characterization of *Mtb*MaK using size exclusion chromatography and sedimentation velocity experiments suggests that *Mtb*MaK exists as a mixture of monomers, trimers and tetramers in solution[7]. Notably, MaK forms a complex with TreS *in vitro* and *in vivo*[7]. This is not surprising partly

due to the fact that the *mak* gene is usually linked with the *treS* gene and in some micro-organisms like *Psuedomonas entamophila* and *Rubrobacter xylanophilus* the two genes are fused into a single gene[16]. Four units of *Mtb*MaK associate with a tetramer of TreS to form a hetero-octameric complex[7]. More importantly, the formation of the complex enhances the activity of *Mtb*MaK by 3-folds[7]. Although the crystal structure of the tetramer of TreS is known[7,17,18], the molecular basis for the enhancement of activity of MaK by TreS is currently not understood. In addition, the structural basis for conversion of maltose to maltose-1-phosphate is not known.

Here, we describe the crystal structures of the *apo*- and maltose-bound *Mtb*MaK. Unexpectedly, the structures reveal a unique N-lobe that has not been observed before in structures of other kinases. MaK uses homotypic dimerization to assemble the active sites for catalysis. Maltose binds in a spherical cavity of the C-lobe with the O1′ oxygen atom oriented towards the catalytic D322. Instead of a glycine-rich loop, MaK uses a ^{142}EQSNXS^{147} motif conserved across MaK homologues to catalyse the phosphorylation of maltose. These structural variations and a putative role for the signal independent homotypic dimerization of MaK in scaffolding are discussed.

Results

Structure of *Mtb*MaK reveals a unique N-lobe. Structure of the *apo*-enzyme was solved by single-wavelength anomalous diffraction (SAD) method using crystals derivatized with $HgCl_2$ (Table 1). The final model consists of amino acids 5–455, 154 water molecules, 1 SO_4^{2-} ion and 3 mercury atoms refined to 2.4 Å resolution with an R factor of 0.207 and a free R value of 0.244. The structure of maltose bound *Mtb*MaK was solved by molecular replacement using the structure of the unliganded enzyme as a search template and consists of 450 amino acids, 1 molecule of maltose, 2 molecules of Bis-Tris, and 1 SO_4^{2-} ion refined to 2.9 Å resolution with an R factor of 0.239 and a free R value of 0.281.

The overall structure of *Mtb*MaK can be divided into an N-lobe primarily composed of β strands and a C-lobe made up of mostly α-helices (Fig. 1a). Unlike typical kinases, the N-lobe of MaK contains not one, but two β-sheets. Each sheet encloses an α helix and a short 3_{10} helix. The sheet proximal to the N-terminus of *Mtb*MaK is highly twisted and is formed by three long (β1–β3) and two short (β4 and β5) strands running anti-parallel to each other. In contrast, the second sheet is made of seven anti-parallel strands (β6–β7, β9–β11, β8′–β9′), that includes two strands (β8′–β9′) contributed by an adjacent monomer. In exchange, strands β8 and β9 of this sheet protrude out of the protein, insert themselves into the adjacent monomer and make up two strands of the corresponding sheet in the adjacent monomer. This results in the formation of an unusual homotypic dimer of *Mtb*MaK (Fig. 1b). The unique arrangement of the structural elements of the N-lobe of *Mtb*MaK constitutes a new fold. In addition, the signal independent homotypic dimerization of *Mtb*MaK involving mutual exchange of structural elements from the N-lobe, has not been observed before for any kinases. In contrast to the N-lobe that is very different from the N-lobes of most kinases, the overall structure of the C-lobe of *Mtb*MaK resembles the C-lobes found in most kinases[19].

Comparison with homologous kinases. 3D structure alignment-based similarity search using Dali[20] retrieved methylthioribose (MTR) kinase (PDB code 2PUN; 13% sequence identity, r.m.s. deviation of 3.9 Å for 249 overlapping C_α atoms; Fig. 1c)[21] and choline kinase (PDB code 1NW1; 16% sequence identity, r.m.s. deviation of 3.6 Å for 228 overlapping C_α atoms)[22] as structural matches. Other significant hits included aminoglycoside phosphotransferases

Table 1 | Data collection and refinement statistics

Dataset	Hg derivative (peak)	*Mtb*MaK (native)	*Mtb*MaK-maltose
Data collection			
Space group	$P6_522$	$P6_522$	$P6_522$
Unit-cell parameters			
a, b, c (Å)	a = b = 97.3, c = 464.0	a = b = 96.6, c = 459.7	a = b = 96.7, c = 461.3
α, β, γ (°)	α = β = 90.0, γ = 120.0	α = β = 90.0, γ = 120.0	α = β = 90.0, γ = 120.0
Wavelength (Å)	1.0052	1.0000	1.0000
Resolution (Å)	50.00–3.20 (3.31–3.20)[a]	50.00–2.40 (2.49–2.40)	50.00–2.90 (3.00–2.90)
Total observation	282,429	910,859	195,018
Unique reflections	22,780	51,256	29,699
Data completeness (%)	99.4 (99.1)	100.0 (100.0)	99.7 (99.8)
R_{merge}[b]	0.138 (0.753)	0.082 (0.548)	0.062 (0.549)
$I/\delta(I)$	16.2 (3.6)	35.3 (4.2)	25.5 (3.3)
Redundancy	12.4 (12.5)	17.6 (9.2)	6.6 (6.7)
Refinement			
Resolution (Å)		38.31–2.40	44.60–2.90
Reflections used		50,237	28,826
R_{work}/R_{free}[c]		0.207/0.244	0.239/0.281
No. of non-H atoms			
Chain A/B		3,349/3,490	3,349/3,490
Ligands and Water		159	56
Average B factor (Å^2)			
Chain A/B		67.4/63.1	116.0/116.9
Ligands and Water		50.5	113.2
R.m.s. deviation			
Bond lengths (Å)		0.017	0.021
Bond angles (°)		1.249	1.699
Ramachandran plot (%)			
Favored		96.9	96.8
Allowed		3.1	3.2
Outliers		0	0

[a]Values in parentheses are for the highest resolution shell.

[b]$R_{merge} = \Sigma_h\Sigma_i \left| I_{ih} - \langle I_h \rangle \right| / \Sigma_h\Sigma_i \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean intensity of the observations of I_{ih} of reflection h.

[c]$R_{work} = \Sigma_h \left| F_o - F_c \right| / \Sigma_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h. R_{free} is mathematically equivalent to R_{work}, but was measured over 5% of the data.

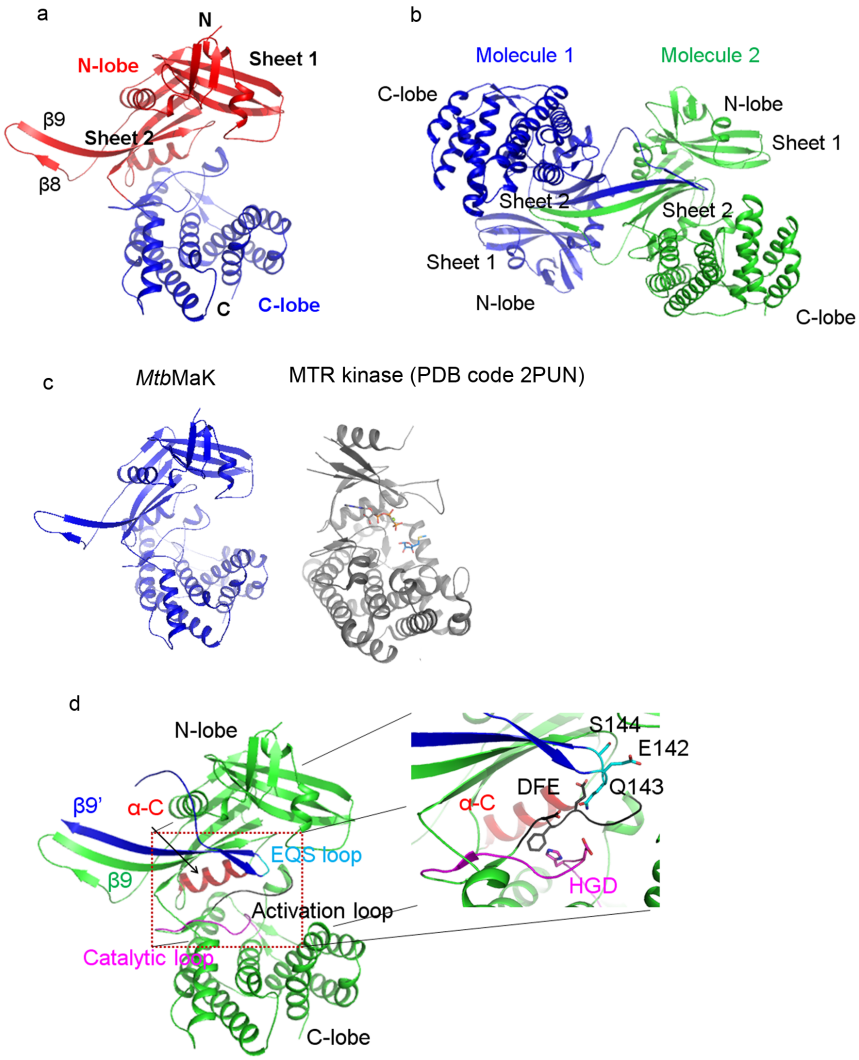

Figure 1 | Structure of *Mtb*MaK. (a). Cartoon representation of a monomer of *Mtb*MaK. The N-lobe contains two β-sheets with strands β8 and β9 protruding out of the protein. The N- and C-terminals of the protein are marked as N and C, respectively. (b). Cartoon representation of a homotypic dimer of *Mtb*MaK. Strands β8 and β9 from each monomer within a dimer mutually insert into each other's N-lobe to assemble sheet 2. (c). Structural homologues of *Mtb*MaK. A Dali analysis of *Mtb*MaK (left) retrieved low structural matches like methylthioribose (MTR) kinase from *Bacillus subtilis* (right; PDB code 2PUN) that could not be superimposed over *Mtb*MaK. (d). Location of conserved motifs. The α-C helix (shown in red), activation loop (black color) and catalytic loop (magenta color) of *Mtb*MaK are located at the junction of N- and C-lobes. A '^{142}EQS^{144}' loop (cyan color) connecting two strands of an adjacent monomer (blue color) contributes residues critical for catalysis. DFE and HGD motifs of *Mtb*MaK as well as residues of '^{142}EQS^{144}' loop are shown as sticks (inset).

and many eukaryotic protein kinases such as cell division protein kinase, casein kinase, proto-oncogene serine/threonine-protein kinase and cAMP-dependent protein kinase (PKA).

Since protein kinases have been characterized extensively and also because *Mtb*MaK shares similarities with eukaryotic protein kinases, key features of *Mtb*MaK are discussed in context with terminologies used for typical protein kinases. Comparisons with prokaryotic structural homologues like MTR and choline kinases bring out remarkable differences; illustrating *Mtb*MaK has evolved specific features for its function. A common conserved feature found in most kinases is the presence of a glycine-rich loop (GxGxxG) between strands β1 and β2. Variations in composition of this loop that is involved in positioning of phosphate groups of ATP have been observed previously [2,12]. The structure of *Mtb*MaK unveils a new variation, an ^{142}EQS^{144} sequence, at a structurally equivalent position (Fig. 1d). Interestingly, MTR kinase has a GxGNxN motif for binding ATP[21] and choline kinase shows the presence of a di-glycine within the β1-β2 loop[22]. A lysine from β3 strand of the N-lobe usually makes contact with α and β phosphates of ATP. K157 from strand β9 of *Mtb*MaK and K61 from strand β3 of MTR kinase are located in equivalent positions. In choline kinase, R111 replaces the lysine at this position. Instead of a glutamate, D169 from the αC helix of *Mtb*MaK makes a salt bridge with K157, analogous to the salt bridge for the formation of an activated "αC-in" conformation of protein kinases. In case of MTR kinase, E84 located on αC helix makes a typical salt bridge with K61. The corresponding glutamate E125 is dis-ordered in the structure of choline kinase. Absence of salt bridge signifies an inactive conformation. This is in agreement with the fact that choline kinase was crystallized in presence of Ca^{2+} ions that are known to inactivate the kinase[12]. The active site of kinases usually contain two aspartate residues; one each located on the catalytic and activation loops. D322 from the ^{320}HGD^{322} motif of *Mtb*MaK is within hydrogen bonding distance of O1′ hydroxyl oxygen of maltose and can function as the key catalytic residue. D233 of MTR kinase located within a similar HGD motif makes contact with O1′ hydroxyl oxygen of the substrate. Intriguingly, choline kinase has a

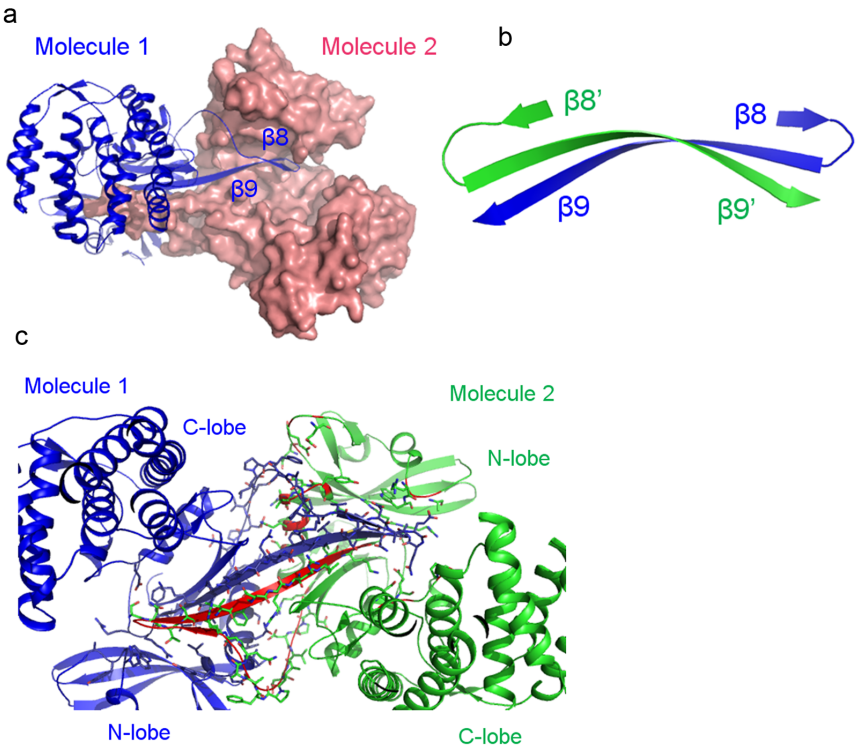

Figure 2 | Homotypic dimerization of *Mtb*MaK. (a). A dimer of *Mtb*MaK where one molecule is depicted as cartoon (blue color), while the second molecule is shown in a surface representation (salmon color). Mutual insertion of strands β8 and β9 results in the formation of a tight dimer. (b). Strand β9 of each monomer (green and blue) run anti-parallel with the backbone atoms interacting along the entire length of the strand. (c). Dimer interface of *Mtb*MaK. Interfacing residues are shown in sticks. Region of molecule 2 participating in dimerization is marked with red color on the cartoon. Side chains as well as backbone atoms participate in dimerization.

HND motif with D255 occupying a catalytically equivalent position. The second aspartate of the active site is part of the activation loop that positions the Mg^{2+} ions for catalysis. D339 of *Mtb*MaK from the ^{339}DFE^{341} motif, D250 of MTR kinase and D301 of choline kinase located within a DXE motif are well positioned to perform this role. Thus, although *Mtb*MaK shares many catalytic features with other kinases, key differences like presence of the 'EQS' loop, which is surprisingly widely distributed in MaK homologues, suggests function-specific alterations of the kinase.

***Mtb*MaK exhibits a unique mode of dimerization.** Analysis of the asymmetric unit of the crystallized *Mtb*MaK revealed the presence of two molecules of the protein. These two molecules of *Mtb*MaK form a dimer *via* homotypic interactions that involve backbone atoms as well as side chains. Strands β8 and β9 from one monomer protrude out and insert into the active site of another monomer (Fig. 2a). Amino acids N145-F159 making up strand β9 of the monomers within a dimer run anti-parallel with the backbone atoms from both the strands interacting along the entire length of the strands (Fig. 2b). Additional intermolecular interactions are observed between side chains of residues from the ^{142}EQS^{144} loop and the loop connecting helix α1 with strand β1. Interestingly, intermolecular interactions between the ^{142}EQS^{144} loop and the "activation loop" are also observed. Further, the loop connecting strand β7 with β8 interacts with residues connecting strands β3 with β4 and helix α2 (Fig. 2c). Details of amino acids involved in inter-molecular interactions are listed in Table S1. PISA[23] analysis revealed that dimerization of *Mtb*MaK buries 2,575 Å^2 of solvent accessible surface area. Intriguingly, structural matches like MTR kinase, choline kinase and aminoglycoside phosphotransferase type IIIa (APH IIIa) retrieved by Dali analysis for *Mtb*MaK, are also known to form dimers[21,22,24]. However, dimerization of these kinases involves participation of C-lobe of one of the monomers. On the other hand, crystal structure of the kinase domain of a transmembrane Thr/Ser kinase, PknB (PDB code 1MRU)[25], does show the N-lobe forming intermolecular interactions. But this mode of dimerization mediated by the N-lobe is very different from that of MaK. While dimerization involves surface residues in PknB, structural elements are mutually exchanged between the N-lobes of two monomers of MaK. As a result, dimerization of *Mtb*MaK buries a much larger solvent accessible surface area (2,575 Å^2) when compared to that buried during dimerization of PknB (1,045 Å^2). Dimerization involving mutual exchange of structural elements has been seen previously in human ChPK2 (Supplementary Fig. S2)[26]. However, ChPK2 dimerization involves the transactivation loop (T-loop) of the C-lobe and not the N-lobe. Thus, the mode of dimerization of *Mtb*MaK is unique. Extensive inter-subunit interactions result in the formation of a tight dimer of *Mtb*MaK. This dimerization of *Mtb*MaK is signal independent and involves homotypic interactions.

Maltose-binding site of *Mtb*MaK. To gain insights into the nature of maltose binding site, we solved the structure of *Mtb*MaK in complex with maltose. As observed for the unliganded enzyme, *Mtb*MaK bound with maltose crystallized as a dimer. Electron density for maltose in Chain B was unambiguous (Fig. 3a). Maltose binding induces conformational changes around the substrate binding site. However, no movement of the lobes is observed. Side chains of amino acids like Y370, Y429, E430, Y433 and R438 move inwards to interact with maltose (Fig. 3b). Detailed analysis of the substrate-binding pocket revealed that 11 residues, W222, D322, H324, P344, S367, Y370, K426, Y429, E430, Y433 and R438, are interacting with maltose (Fig. 3c–e and Supplementary Table S2). These interactions fix the orientation of maltose for catalysis. Notably, the OD2 carboxyl oxygen of the catalytic D322 is observed forming a hydrogen bond with the O1′ atom of maltose. *Mtb*MaK exhibited a *Km* value of

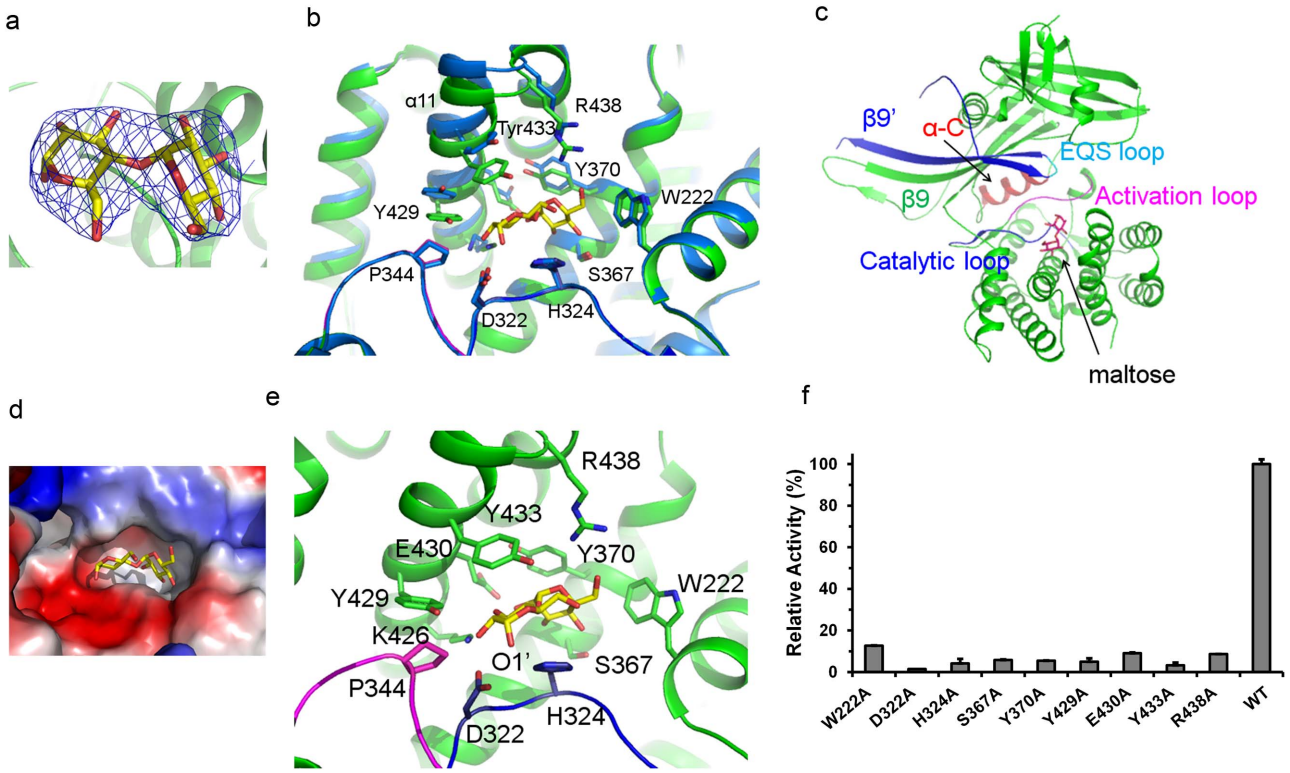

Figure 3 | Maltose binding site of *Mtb*MaK. (a). 2Fo-Fc electron density for maltose contoured at 1.5σ is shown. (b). Conformational changes upon binding of maltose. The structure of maltose bound *Mtb*MaK (green color) was superimposed over the unliganded structure (blue color). Main and side chains of several residues move as a result of substrate binding. (c). Maltose binds in the C-lobe and is in proximity to the conserved motifs essential for catalysis. (d). A surface electrostatic potential representation of the region around the maltose binding site. Blue represents positive potential; red, negative potential. (e). Residues interacting with maltose. Residues from the catalytic loop interacting with maltose are shown as blue sticks, while P344 from the activation loop is shown as magenta sticks. Maltose is shown as sticks in panels A-E. (f). Alanine scanning mutagenesis of amino acids interacting with maltose. Relative activity of mutants (%) with respect to the wild type is plotted as a bar graph. Error bars represent s.d. (n = 3).

2.49 mM for maltose (Fig. S7). This value is similar to the previously reported value for *M. bovis* BCG MaK (2.52 mM)[16]. To verify the role of the amino acids interacting with maltose in catalysis, we performed alanine-scanning mutagenesis and tested the ability of the mutants to phosphorylate maltose. Except for the P344A and K426A mutants, all other mutants could be expressed as soluble protein in *E. coli*. All nine mutants exhibited a dramatic loss in activity (activity less than 13% of the wild type) (Fig. 3f and Supplementary Table S4). Taken together, the structural observations coupled with results of mutagenesis studies suggest that Y222, D322, H324, S367, Y370, Y429, E430, Y433 and R438 are crucial for phosphorylation of maltose. These amino acids are conserved across MaK homologues from different sources (Supplementary Fig. S3).

The structure of the binary complex of *Mtb*MaK with maltose helps explain the structural basis for α maltose-specific (Supplementary Fig. S4) phosphorylation by MaK. The aromatic ring of Y433 stacks above the glycosyl moiety with the reducing end (Fig. 3e). Specifically, the C1 atom of the sugar is 3.5 Å away from the aromatic ring. In its β form, the O1 atom of maltose would either sterically clash with the aromatic ring of Y433 or result in non-productive binding. This is because stacking of Y433 with the glycosyl moiety brings catalytically critical residues in proximity of maltose (Fig. 3b). A β form of the maltose would not permit this stacking. In addition, the catalytic machinery is positioned for phosphorylation of the O1′ atom of the α isomer. The O1′ atom of maltose's β form would be positioned far away from D322 to permit phosphorylation at this position.

Nucleotide-binding site of *Mtb*MaK. Although there is no significant sequence conservation between *Mtb*MaK and its homologues over the entire length of the protein, some conservation around the ATP-binding site is observed (Fig. 4a). To gain insights into the ATP-binding site of *Mtb*MaK, the nucleotide-binding site of aminoglycoside phosphotransferase (APH; PDB code 1J7U) was superimposed over the structure of *Mtb*MaK. This region of APH superimposed over an equivalent region of *Mtb*MaK with an *r.m.s.d.* of 2.6 Å between the C_α atoms of 90 matching residues. Using such an approach, AMPPNP-Mg of APH could be docked into the nucleotide-binding site of *Mtb*MaK (Fig. 4b) and a putative mode of ATP binding by *Mtb*MaK could be inferred. The mode of ATP binding was verified by performing alanine-scanning mutagenesis of residues of *Mtb*MaK interacting with the docked AMPPNP-Mg moiety (Supplemental text; Fig. 4c and Supplementary Table S4). The analysis revealed that amino acids like E142, Q143 and S144 from the "EQS" loop that is conserved across all MaK homologues (Supplementary Fig. S3) are important for binding ATP.

Discussion

Our structural studies on *Mtb*MaK unveil several new, previously unknown features of maltose kinases. Many kinases dimerize upon receiving a signal and activate cellular signaling pathways[27]. Such dimerization usually involves the C-lobe[28]. In contrast, *Mtb*MaK undergoes signal-independent homotypic dimerization mediated by the N-lobe. Dimerization probably offers a means for *Mtb*MaK to form a scaffold that could readily hetero-oligomerize with the

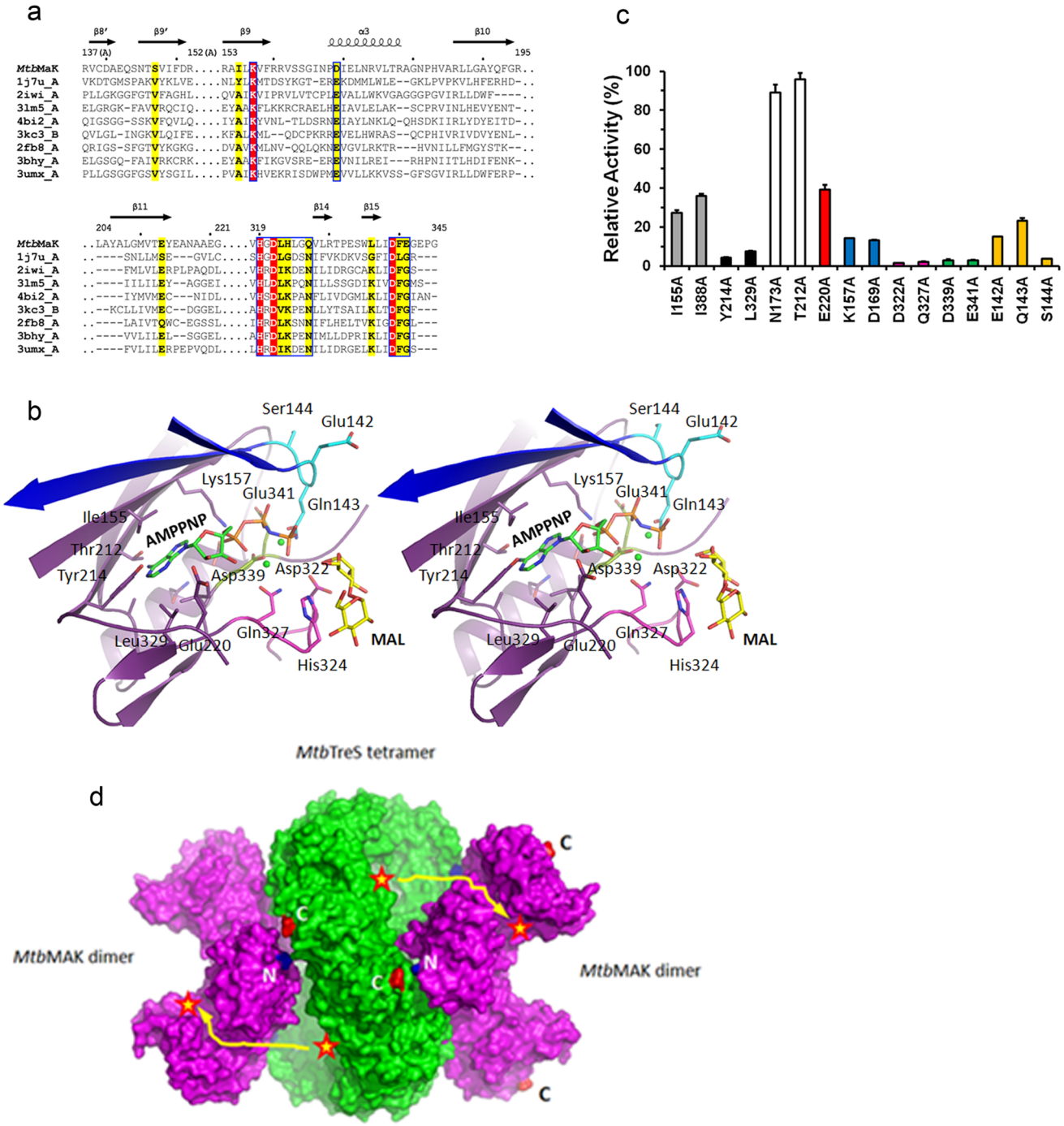

Figure 4 | Nucleotide binding site of *Mtb*MaK. (a). Sequence alignment of nucleotide binding site (NBS) of *Mtb*MaK and its homologues. SSE of *Mtb*MaK are labeled on top of the aligned sequences. Identical residues are highlighted in red, and other conserved residues are highlighted in yellow. (b). Stereo view of AMPPNP-Mg (green sticks and spheres, PDB code 1J7U) superimposed on the NBS of *Mtb*MaK. The C atoms of '[142]EQS[144]' loop, HGD motif and DFE motif are colored in cyan, magenta and light green, respectively. Strands β8′ and β9′ from the other subunit of *Mtb*MaK are colored in blue. The side chains of residues potentially interacting with AMPPNP-Mg are shown as sticks. The O and N atoms are colored in red and blue. (c). Mutagenesis of residues from the NBS. A bar graph of relative activity (%) of mutants compared to the wild type enzyme is shown. Error bars represent s.d. (n = 3). (d). Model of hetero-octameric complex of TreS with MaK. Potential interaction of a tetramer of TreS (green color, PDB code 4LXF) with two dimers of *Mtb*MaK (purple color) is shown. Active sites are marked with a red star; putative path of product marked in yellow.

upstream enzyme, TreS, of the TreS-MaK-GlgE-GlgB pathway (Supplementary Fig. S1). The premises for such a scaffolding role for *Mtb*MaK can be found in the hetero-octameric complex of TreS-MaK reported previously[7]. Further support for the scaffolding role of MaK is provided by the fact that the *mak* gene is usually located immediately downstream of the *treS* gene and in some micro-organisms the two genes are fused into a single gene[16]. Using shape complementarity, a model for interaction of MaK with TreS could be built with the N-terminus of MaK placed in proximity of the C-terminus of TreS, depicting the closely linked nature and order of the genes. In this model, a dimer of MaK fits on each face of a tetramer of TreS (PDB code 4LXF; Fig. 4d). Such an arrangement of

TreS-MaK subunits aligns the active sites of the two proteins in a manner that facilitates the transfer of the product of TreS, maltose, into the active site of MaK for phosphorylation (Fig. 4d).

Scaffolding roles for kinases have been demonstrated previously. For example, the integrin-linked kinase (ILK) functions exclusively as a scaffolding kinase. ILK binds CH2 domain of α-parvin and integrin tails during recruitment of focal adhesion proteins[29]. However, ILK cannot catalyze phospho-transfers due to a defunct active site. In contrast, *Mtb*MaK is active and readily phosphorylates α-maltose. Since maltose-1-phosphate is toxic to mycobacterium, it needs to be converted into other metabolites quickly. This could be accomplished by assembling a super complex of TreS-MaK-GlgE as proposed previously[7]. Using shape complementarity a dimer of GlgE could be possibly fit on the un-occupied face of the MaK dimer (Fig. 4d) giving rise to a super complex in the order GlgE dimer – MaK dimer – Tres tetramer – MaK dimer – GlgE dimer. Thus, MaK kinase could possibly serve as a scaffold for the assembly of a super complex. However, biochemical evidence for a TreS-MaK-GlgE complex has not come forth as yet. Results of our preliminary studies on ability of *Mtb*MaK to form a complex with *Mtb*GlgE using GST-pulldown assays and *in vitro* binding studies of tag-less protein using Biacore suggests that unliganded recombinant proteins do not form a complex. Perhaps, additional factors like ligands or TreS are required for the formation of the complex.

The dimer of *Mtb*MaK observed in the crystal structure probably represents only one of the conformations assumed by the protein. *Mtb*MaK eluted in two peaks during size exclusion chromatography (SEC) runs. Subsequent HPLC analysis revealed a similar pattern (Supplementary Fig. S5A). The retention time of the larger peak suggested presence of a dimer of MaK, while the shorter peak consisted of monomers of MaK. Protein from both the peaks was active when tested for kinase activity (Supplementary Fig. S5B). Analysis of symmetry mates revealed that *Mtb*MaK could possibly form a tetramer (Supplementary Fig. S6). PISA analysis suggested a Complexation Significance Score (CSS) of 0.769 for the tetrameric assembly, indicating the possibility of *Mtb*MaK assuming a tetrameric biological assembly. These results are consistent with previously reported ability of *Mtb*MaK to exist in different oligomeric states in solution[7].

The structures of *Mtb*MaK reported here could be used as guides to develop inhibitors that could potentially aid treatment of TB. Analogues of α-maltose and maltose-1-phosphate could serve as scaffolds that could be re-designed to increase potency and other drug-like properties. Further, the use of MaK inhibitors could be extended to control other pathogens like *Pseudomonas spp*. that harbor the conserved *mak* gene. In this context, lack of *mak* gene in mammalian systems could prove to be beneficial for achieving greater specificity.

Methods

Molecular cloning. The gene encoding the full length *Mtb*MaK (Rv0127) was amplified by polymerase chain reaction (PCR) from *Mycobacterium tuberculosis* H37Rv genomic DNA. The product was sub-cloned into expression vector pGEX-6p-1 (GE Lifesciences) after digestion with *Bam*HI and *Xho*I (TaKaRa). Site-directed mutagenesis was performed by PCR using native *Mtb*MaK expression plasmid as a template. All mutants were constructed as per manufacturer's instructions. Clones were verified by sequencing the entire gene. Recombinant plasmid containing either native *mak* gene or the mutants was then transformed into *E. coli* BL21 (DE3) for protein expression. Primers used in molecular cloning are listed in Supplementary Table S3.

Protein expression and purification. Cells containing recombinant plasmid were grown in Luria Bertani medium at 37°C until OD_{600nm} reached 0.8. After cooling the culture, GST-tagged *Mtb*MaK protein expression was induced at 16°C by addition of 0.5 mM Isopropyl β-D-thiogalactoside (IPTG). After 20 h of induction, cells were harvested by centrifugation and re-suspended in 1 × PBS, pH 7.4. *Mtb*MaK was recovered by first lysing the cells using sonication and clarifying the lysate by centrifugation at 30,000 *g* for 30 min to remove cell debris. The supernatant containing soluble recombinant protein was subjected to GST-affinity chromatography. Glutathione Sepharose 4B beads mixed with the supernatant were washed with buffer to remove non-specifically bound proteins. To remove the GST tag, PreScission protease (GE Lifesciences) was added and the mixture incubated overnight at 4°C. Tagless *Mtb*MaK protein was eluted from the column by gravity, concentrated and injected on a Superdex 200 10/300 GL gel filtration column equilibrated with a buffer containing 20 mM Tris-HCl, pH 8.0, and 150 mM NaCl. Fractions containing the dimer were pooled and concentrated to 10 mg/mL for crystallization. Mutant proteins were expressed and purified in a similar way.

Crystallization and structure determination. Crystallization screening experiments were first carried out at 16°C by vapor diffusion method in 96-well plates using a Mosquito liquid handling system (TTP Labtech). Each crystallization drop contained 200 nL of *Mtb*MaK protein (10 mg/mL) mixed with 200 nL of reservoir solution. Commercially available sparse matrix screens were used for screening crystallization conditions. Crystals grew in a solution containing 1 M $(NH_4)_2SO_4$, 0.1 M Bis-Tris, pH 5.5, and 1% (w/v) PEG3350. Crystals were optimized by varying the concentration of $(NH_4)_2SO_4$ and pH. Optimized crystals were soaked in the reservoir solution supplemented with 20% (v/v) glycerol for 10 s and flash frozen in liquid nitrogen. X-ray diffraction data were collected at 100 K on beamline BL17U at SSRF (Shanghai, China) and processed using HKL2000 (HKL Research, Inc.)[30]. Mercury derivatives were prepared by soaking the native crystals in the reservoir solution supplemented with 1 mM $HgCl_2$ for 2 h prior to cryoprotection. Anomalous peak data were collected on beamline BL41XU at Spring-8 (Hyogo, Japan) and processed to 3.2 Å resolution. The structure of *Mtb*MaK was determined by the single-wavelength anomalous diffraction (SAD) method, using the peak data. All the 3 potential mercury atoms covalently linked to three cysteine residues were located using SHELXD[31], and the initial phases were calculated using PHENIX[32]. The structural model was initially built according to the electron density map and then refined using the native data collected to 2.4 Å resolution. Manual adjustments to the model were made using COOT[33] and refined using PHENIX. The *Mtb*MaK-maltose complex was prepared by soaking *Mtb*MaK crystals in the reservoir solution supplemented with 10 mM maltose (Sigma-Aldrich) for 1 h. The diffraction data for the complex were collected on beamline BL5A at KEK (Tsukuba, Japan) and processed to 2.9 Å resolution. The *Mtb*MaK-maltose complex structure was solved by the molecular replacement (MR) method with Phaser program[34] in CCP4[35], using the native structure as a search template. The complex structure was refined using the native structure as a reference model in order to restrain the dihedral angles. The final structural models of native *Mtb*MaK and *Mtb*MaK-maltose complex have R_{work}/R_{free} values of 0.207/0.244 and 0.239/0.281, respectively. Data collection and refinement statistics are listed in Table 1. Structural figures were prepared using PyMol (DeLano Scientific).

Enzyme activity assay. Kinase activity was estimated by a linked spectrophotometric assay that uses pyruvate kinase and L-lactate dehydrogenase as auxiliary enzymes to monitor the conversion of ATP to ADP with an observable decrease in the absorption of NADH at 340 nm[36]. 100 µL of the reaction mixture contained 2 µg *Mtb*MaK, 10 mM $MgCl_2$, 1.5 mM ATP, and 5 mM maltose suspended in a buffer made up of 20 mM Tris-HCl, pH 8.0, 150 mM NaCl. 3 U of pyruvate kinase, 3 U of L-lactate dehydrogenase, 0.3 mM NADH and 2.5 mM phosphoenolpyruvate (all from Sigma-Aldrich), were added and the mixture incubated at 37°C for 30 min. The change in A_{340nm} value was measured using a VarioshanFlash spectrophotometer (Thermo Scientific) over the course of the incubation period. The velocity of formation of the product ADP was calculated. The *Km* value of *Mtb*MaK for maltose was determined from the Lineweaver-Burk plot constructed by measuring enzyme activity using different concentrations of the substrate.

1. Iturriaga, G., Suarez, R. & Nova-Franco, B. Trehalose metabolism: from osmoprotection to signaling. *Int J Mol Sci* **10**, 3793–3810 (2009).
2. De Smet, K. A., Weston, A., Brown, I. N., Young, D. B. & Robertson, B. D. Three pathways for trehalose biosynthesis in mycobacteria. *Microbiology* **146** (Pt1), 199–208 (2000).
3. Elbein, A. D., Pan, Y. T., Pastuszak, I. & Carroll, D. New insights on trehalose: a multifunctional molecule. *Glycobiology* **13**, 17R–27R (2003).
4. Winder, F. G., Tighe, J. J. & Brennan, P. J. Turnover of acylglucose, acyltrehalose and free trehalose during growth of Mycobacterium smegmatis on glucose. *J Gen Microbiol* **73**, 539–546 (1972).
5. Mendes, V., Maranha, A., Alarico, S. & Empadinhas, N. Biosynthesis of mycobacterial methylglucose lipopolysaccharides. *Nat Prod Rep* **29**, 834–844 (2012).
6. Sambou, T. *et al.* Capsular glucan and intracellular glycogen of Mycobacterium tuberculosis: biosynthesis and impact on the persistence in mice. *Mol Microbiol* **70**, 762–774 (2008).
7. Roy, R. *et al.* Synthesis of alpha-Glucan in Mycobacteria Involves a Hetero-octameric Complex of Trehalose Synthase TreS and Maltokinase Pep2. *ACS Chem Biol* **8**, 2245–2255 (2013).
8. Kalscheuer, R. *et al.* Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an alpha-glucan pathway. *Nat Chem Biol* **6**, 376–384 (2010).
9. Pan, Y. T. *et al.* Trehalose synthase converts glycogen to trehalose. *FEBS J* **275**, 3408–3420 (2008).
10. Elbein, A. D., Pastuszak, I., Tackett, A. J., Wilson, T. & Pan, Y. T. Last step in the conversion of trehalose to glycogen: a mycobacterial enzyme that transfers

maltose from maltose 1-phosphate to glycogen. *J Biol Chem* **285**, 9803–9812 (2010).
11. Pal, K. *et al.* Crystal structure of full-length Mycobacterium tuberculosis H37Rv glycogen branching enzyme: insights of N-terminal beta-sandwich in substrate specificity and enzymatic activity. *J Biol Chem* **285**, 20897–20903 (2010).
12. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77–84 (2003).
13. Griffin, J. E. *et al.* High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* **7**, e1002251 (2011).
14. Sani, M. *et al.* Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog* **6**, e1000794 (2010).
15. Niehues, B. *et al.* Isolation and characterization of maltokinase (ATP:maltose 1-phosphotransferase) from Actinoplanes missouriensis. *Arch Microbiol* **180**, 233–239 (2003).
16. Mendes, V., Maranha, A., Lamosa, P., da Costa, M. S. & Empadinhas, N. Biochemical characterization of the maltokinase from Mycobacterium bovis BCG. *BMC Biochem* **11**, 21 (2010).
17. Caner, S. *et al.* The structure of the Mycobacterium smegmatis trehalose synthase reveals an unusual active site configuration and acarbose-binding mode. *Glycobiology* **23**, 1075–1083 (2013).
18. Ravaud, S. *et al.* Trehalulose synthase native and carbohydrate complexed structures provide insights into sucrose isomerization. *J Biol Chem* **282**, 28126–28136 (2007).
19. Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci* **36**, 65–77 (2011).
20. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–549 (2010).
21. Ku, S. Y. *et al.* Structures of 5-methylthioribose kinase reveal substrate specificity and unusual mode of nucleotide binding. *J Biol Chem* **282**, 22195–22206 (2007).
22. Peisach, D., Gee, P., Kent, C. & Xu, Z. The crystal structure of choline kinase reveals a eukaryotic protein kinase fold. *Structure* **11**, 703–713 (2003).
23. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797 (2007).
24. Burk, D. L., Hon, W. C., Leung, A. K. & Berghuis, A. M. Structural analyses of nucleotide binding to an aminoglycoside phosphotransferase. *Biochemistry* **40**, 8756–8764 (2001).
25. Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M. & Alber, T. Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. *Nat Struct Biol* **10**, 168–174 (2003).
26. Oliver, A. W. *et al.* Trans-activation of the DNA-damage signalling protein kinase Chk2 by T-loop exchange. *EMBO J* **25**, 3179–3190 (2006).
27. Lassila, J. K., Zalatan, J. G. & Herschlag, D. Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu Rev Biochem* **80**, 669–702 (2011).
28. Bartek, J. & Lukas, J. DNA damage checkpoints: from initiation to recovery or adaptation. *Curr Opin Cell Biol* **19**, 238–245 (2007).
29. Fukuda, K., Gupta, S., Chen, K., Wu, C. & Qin, J. The pseudoactive site of ILK is essential for its binding to alpha-Parvin and localization to focal adhesions. *Mol Cell* **36**, 819–830 (2009).
30. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326 (1997).
31. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**, 1772–1779 (2002).
32. Echols, N. *et al.* Graphical tools for macromolecular crystallography in PHENIX. *J Appl Crystallogr* **45**, 581–586 (2012).
33. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).
34. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
35. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* **59**, 1131–1137 (2003).
36. Shan, S. *et al.* Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from Mycobacterium tuberculosis. *FASEB J* **25**, 1577–1584 (2011).

Acknowledgments

We would like to thank staff members of the Core Facility for Protein Research of the Institute of Biophysics, CAS; especially Yi Han, Ya Wang and Xiaoxia Yu for their technical support. This work was supported by grants from the State Key Development Program for Basic Research of the Ministry of Science and Technology of China (973 Project Grant Nos. 2014CB542800, 2011CB915501 and 2011CB910304), National Science Foundation Grant 813300237, and National Infectious Disease Funding (Grant No. 2012ZX10004701).

Author contributions

J.L. and X.G. performed experiments, interpreted results and wrote the manuscript. N.S. analyzed data and participated in drafting the article for important intellectual content. W.C., Y.D. and X.X. collected crystallographic data and analyzed it. Z.R. and X.L. initiated the study, co-designed experiments, interpreted data and wrote the manuscript. All authors have contributed to, seen and approved the manuscript.

Additional information

Accession codes: The co-ordinates and structure factor files for *apo-* and maltose-bound MaK have been deposited in the PDB under accession codes 4O7O and 4O7P, respectively.

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

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

How to cite this article: Li, J. *et al.* Homotypic dimerization of a maltose kinase for molecular scaffolding. *Sci. Rep.* **4**, 6418; DOI:10.1038/srep06418 (2014).