Biochemistry and Biophysics Reports 33 (2023) 101410

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Characterizing a novel CMK-EngA fusion protein from *Bifidobacterium*: Implications for inter-domain regulation

Ammu Suresh [a,b], Balaji Prakash [c,**], Kiran Kulkarni [a,b,*]

[a] CSIR-National Chemical Laboratory, Biochemical Science Division, Dr. Homi Bhabha Road, Pune, 411008, India
[b] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, India
[c] Division of Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Central Campus, Navrangpura, Ahmedabad, 380009, India

ARTICLE INFO

Keywords:
Bifidobacterium longum
EngA
Enzyme kinetics
Gene fusion
ITC

ABSTRACT

EngA is an essential and unique bacterial GTPase involved in ribosome biogenesis. The essentiality and species-specific variations among EngA homologues make the protein a potential target for future drug development. In this aspect, it is important to understand the variations of EngA among probiotic organisms and non-probiotic bacteria to understand species specificity. The search for variations among EngA homologues revealed a unique variant, exclusively found in *Bifidobacterium* and a few *Actinobacteria* species. *Bifidobacterium* possesses a multifunctional fusion protein, wherein EngA is fused with an N-terminal CMK (Cytidylate Monophosphate Kinase) domain. The resulting protein is therefore a large (70kDa size) with 3 consecutive P-loops and a 50 amino acid long linker connecting the EngA and CMK domains. EngA is known to regulate ribosome biogenesis via nucleotide-dependent conformational changes. The additional domain may introduce further intricate regulation in ribosome biogenesis or participate in newer biological processes. This study is the first attempt to characterise this novel class of bacterial EngA found in the Genus of *Bifidobacteria.*

1. Introduction

Ribosome biogenesis is an essential biological process in all living cells. Given its complexity, it is not surprising that it is regulated by myriad proteins, including the EngA family of GTP binding Proteins [1,2]. Knocking off EngA resulted in the accumulation of immature ribosomes and a compromised central protuberance of the large subunit [3]. The study hinted at an important role for EngA in the development of central protuberance of the 50S ribosomal subunit. EngA consists of two tandem canonical GTPase domains (GD1 and GD2) at the N-terminus, followed by a KH domain [4]. Despite exhibiting significant sequence similarity, GD1 and GD2 show different affinities for Guanine nucleotides and also exhibit different hydrolysis rates [5]. Based on the biochemical data it has been suggested that EngA can acquire nine different conformational states depending on the nucleotide bound (or unbound Apo) states of its two G-domains GD1 and GD2 [6]. This conformational flexibility was suggested to be critical in regulating ribosome biogenesis [6]. Mutations in EngA have bactericidal effects on several organisms, including pathogens like *Salmonella typhi* [7]. Sequence analysis of several EngA homologs showed distinct variations among firmicutes and non-firmicutes. Such variations were majorly found in the linker region connecting the GD1 and GD2 domains and also at the C-terminal tail [8].

Emerging antibiotic resistance in pathogens has spurred research toward the development of novel antibiotics. However, recent advancements in microbial research have underlined the role of microbiota on host-pathogen physiology [9–11]. Therefore, it is now imperative that we develop species-specific antibiotics that specifically target the pathogens while protecting the "native" microbiota of the host.

EngA was shown to be a potential drug target against lethal diseases like Tuberculosis [12] and other bacterial diseases. Towards developing species-specific antibiotics, we set out to understand species-specific variations in this class of proteins, as they would provide insights for the design of highly specific and potent antibiotics against pathogens. We began by analysing variations among EngA homologues across the bacterial kingdom. Interestingly, we came across a unique EngA variant, found exclusively in the *Actinomycetes* phylum. Other than the canonical GD1, GD2 and KH domains, this variant of EngA has a fused N-terminal Cytidine Monophosphate Kinase (CMK) domain, which possesses the

* Corresponding author. CSIR-National Chemical Laboratory, Biochemical Science Division, Dr. Homi Bhabha Road, Pune, 411008, India.
** Corresponding author.
E-mail addresses: balaji.prakash@ahduni.edu.in (B. Prakash), ka.kulkarni@ncl.res.in (K. Kulkarni).

https://doi.org/10.1016/j.bbrep.2022.101410
Received 3 October 2022; Received in revised form 1 December 2022; Accepted 7 December 2022
Available online 16 December 2022
2405-5808/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

canonical CMP Kinase motifs. Incidentally, all the *Bifidobacterium* species possess this unique variant of EngA, and hence we have named this BiCMKEngA.

The essentiality of CMK has been debated in literature [13–16]. It is an essential gene in a few bacteria but not in a few others. In eukaryotes, the CMP kinase is a multifunctional enzyme, which can convert not only CMP but also UMP and its deoxy monophosphates to their respective dinucleotide form [17]. In contrast, Bacteria have separate kinases to act on these two nucleotides; CMP kinases to convert CMP and dCMP to their dinucleotide, and Aspartate Kinases to convert UMP to UDP [18, 19]. *Bifidobacterium* too possesses a separate UMP kinase for the conversion of UMP to UDP. Another interesting aspect concerns the phosphate donor in the CMKs. While *E.coli* CMK can use either GTP or ATP as the phosphate donor for CMP conversion [20,21], the *Bacillus Subtilis* CMK cannot use GTP [21]. CMK is also known to exhibit substrate-assisted conformational changes upon CMP binding [17,22].

However, structure of CMK bound with just the phosphate donor is not available. CMK requires CMP (either of the monophosphates) to undergo a conformational change required for ATP /GTP binding. A region termed the lid is floppy and is stabilized upon CMP binding the active site as can be seen in the structure of the CMK homologue from *Thermatoga maritima* [22].

Among bacteria, many CMK homologues are found to be fused with other kinases, but none with the GTPases. BiCMKEngA is the first such example of a fusion of CMK with EngA a dual GTPase domain. Phylum *Actinomycetes* consists of a large variety of species and it is one of the dominant phyla found in the bacterial kingdom. However, the essential GTPase EngA is not fused with CMK in all the species of *Actinomycetes*, indicating a likely evolutionary event that resulted in the fusion or separation of the kinase-GTPase gene during species divergence. *Bifidobacterium* is well studied for its probiotic effects, PAN-Genome and saccharolytic enzymes, but very little is known about the enzymes involved in essential metabolic pathways. Unfortunately, variations in the essential metabolic pathways of probiotic species are still poorly studied. Understanding the variations found in the essential metabolic pathways and the enzymes therein, of probiotic organisms is, therefore, an important step toward developing species-specific antibiotics.

In this work, we report the evolutionary origin of CMK-fused EngA proteins and the biochemical characterisation of BiCMKEngA. We find that the CMK domain of BiCMKEngA has promiscuity for the phosphate donor, i.e. it can utilize both ATP and GTP, which may have functional implications in ribosome maturation. By employing modelling studies, we have delineated the specificity governing factors for the ligand (dCMP, CMP) and cofactors (phosphate donors - ATP, GTP) of the enzyme. We also attempt to address the influence of the CMK domain on the GTPase activity of the *Bifidobacterium* BiCMKEngA and vice versa. Overall, we have characterised a new variant of EngA, a multifunctional BiCMKEngA from *Bifidobacterium longum.*

2. Materials and methods

2.1. Multiple sequence alignment and homology modelling

Protein sequences of EngA and CMKEngA were obtained from UniProtKB [23], and identical sequences were removed using the CDhit Suit (http://weizhong-lab.ucsd.edu/cdhit-web-server/cgi-bin/index.cgi) [24]; we omitted 70% identical sequences. Muscle13.3.8 [25] was used for creating Multiple Sequence Alignments (MSA), and Jalview [26] for viewing the MSA files. Mega X prototype [27] is used for alignment and phylogenetic tree creation. IToL [28] web server is used for viewing and editing phylogenic trees. Synteny analysis of the genes in the bacteria is done using the SynTax web server [29].

The amino acid sequences of the protein were retrieved from UniProtKB and Robetta [30] was used for the construction of the homology model. The obtained model was used for comparison with the available PDB structures of CMK and EngA.

2.2. Cloning expression and purification of BiCMKEngA from Bifidobacterium longum

Gibson assembly cloning method was used for cloning the *Bifidobacterium longum* EngA (BiCMKEngA) and other constructs described here. The gene was cloned from *Bifidobacterium longum* strain NCIM 5684. The primers are obtained from Sigma and the details are in Supplementary Table S1. Cloned plasmids were transformed into BL21 DE3 cells, induced with 0.2mM of IPTG at 18 °C temperature overnight. The pellets were harvested using Beckman Avanti-JXN 100 centrifuge, rotor JLA- 8.1 rotor at 4000rpm, for 20 min. The pellet was lysed in buffer A (all buffer compositions are in Supplementary Table S2). The proteins were purified using affinity purification chromatography using a HisTrap 5ml column. Proteins were eluted using imidazole gradient elution at an imidazole concentration of 150mM. The eluted proteins were further concentrated using a Centricon of membrane size of 30kDa for the full-length BiCMKEngA and a size 10KDa for both BiEngA, and BiCMK – the individual domains of the BiCMKEngA protein. The concentrated proteins were further purified using size exclusion chromatography using a Sephadex 200 column. The eluted fractions were pooled concentrated and snap frozen using Liq.N2 and stored at −80 °C deep freezer. The eluted proteins were subjected to MALS using Agilent 1260 Infinity Autosampler with Wyatt Dawn Heleos II Multi-Angle Light Scattering (MALS) detector using a Superdex® 200 Increase 10/300GL column (Supplementary Figs. S5A, 5B, 5C). The proteins were eluted using Tris 10mM, NaCl 300mM (pH −7.5) with 5%glycerol ((v/v) at room temperature during MALS analysis.

2.3. Enzyme kinetics

Kinetics parameters of BiCMKEngA and BiCMK for the conversion of CMP to CDP were found using the NADH reduction assay [31], Briefly, for the assay buffer contains 50 mM Tris pH 7.4, 2 mM $MgCl_2$, 50 mM KCl, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 0.5 mM ATP, 2 units each of lactate dehydrogenase, pyruvate kinase, and various concentrations (0.05mM–0.8mM) of CMP/dCMP is used. The assay mixture along with 1 μM of protein incubated at 25 °C and the decrease in absorbance was recorded at 340 nm. Absorbance were taken immediately without prior incubation since it is a real time assay. NADH coupled reduction assay cannot be used for the Km determination of ATP/GTP due to the allosteric inhibitory effect of ATP on pyruvate kinase. The rate of enzyme activity was calculate using a NADH standard plot, and it was measured for different concentrations of CMP/dCMP. MM plots were plotted using Graphpad prism using the enzyme activity measurements.

The GTPase activity assay was performed using a commercially available calorimetric malachite green assay kit (Enzo Biomol reagent). 10μM GTPase (BiCMKEngA, BiEngA) was incubated with different concentrations of GTP, varying from 0.02mM to 1mM in assay buffer (100 mM Tris-HCl, 20 mM $MgCl_2$ and 400 mM KCl, pH 7.5) in 96 well microplates. Reactions were mixed well and incubated for 30 min at ambient temperature before the addition of the Biomol reagent according to the manufacturer's instructions. Both the calorimetric and fluorometric values are collected using BioteK, cytation5 image reader. The MM kinetic plots were fitted using Graphpad prism9.

2.4. Nucleotide binding studies

To understand the nucleotide-binding affinities of the proteins, Iso-Thermal Titration calorimetry (ITC) experiments were carried out. However, the optimal buffer compositions for ITC were arrived at, based on fluorescently labelled nucleotide assays (Supplementary Fig. S6). Nucleotides (GTP and ATP) labelled with the fluorescent group, N-methyl-3′-O-anthranoyl (mant) i.e. mant-GTP and mant-ATP were used for these assays. All the experiments were performed at room temperature using 2 μM of the respective protein and 0.2μM of the mant-

nucleotide. The samples were excited at 340 nm and fluorescence emission was monitored by recording the spectra from 400 to 500 nm using cytation5 Spectrophotometer (BioteK). Only if mant-nucleotide is bound to the protein enhanced emission at 450nm is observed. All the samples were prepared in 10mM Tris pH8.0 and 300 mM NaCl, 50mM $MgCl_2$ and with or without 1mM CMP. This buffer composition increased fluorescence, which was subsequently chosen for ITC experiments as well. ITC was performed using Malvern ITC peakQ systems keeping 20 or 30 μM protein in the cell and 500μM of non-labelled ATP or GTP in the syringe. The experiment was performed at 25 °C with 19 injections of 2 μl each.

3. Results

3.1. Discovery of a unique CMK-EngA multifunctional fusion protein

In our pursuit to understand EngA across the probiotic organisms, we identified a unique EngA homologue from *Bifidobacterium* that we term BiCMKEngA. This protein was found to be fused with CMP kinase (CMK) at the N-terminus (Fig. 1A). Both CMP kinase (CMK) and EngA domains exhibit modest sequence homology, in the range of 45%, with other bacterial homologues (Supplementary Fig. S1). Although few species from *Actinobacteria* phylum such as *Corynebacterium* also possess a fused CMK-EngA protein, not all species from those genera possess this variant CMK-EngA except the *Bifidobacterium* (BiCMKEngA) (Fig. 1B). Interestingly, the pathogenic species *Mycobacterium* also belongs to the *Actinomycetes* phylum but it does not exhibit the fusion event and possesses CMK and EngA as two separate proteins (Supplementary Figs. S2A and S2B). The phylogenetic proximity of the genes of CMK and EngA was confirmed by synteny analysis of the *Mycobacterium* species, and it was found that both CMK and EngA genes are situated in close proximity in the genome (Supplementary Figs. S2A–S2D). The variations found amongst EngA homologues in this phylum may correlate well with the evolution of a protein associated with pathogenicity, such effects were well studied among yeast [32] and eukaryotes.

3.2. Cloning and heterologous expression of BiCMKEngA

The full-length BiCMKEngA was found to be well expressed and stable in heterologous systems like *E. coli* (Supplementary Fig. S4). To characterise the activity and study intra-domain interactions in the protein, we expressed the EngA and CMK domains (termed BiEngA and

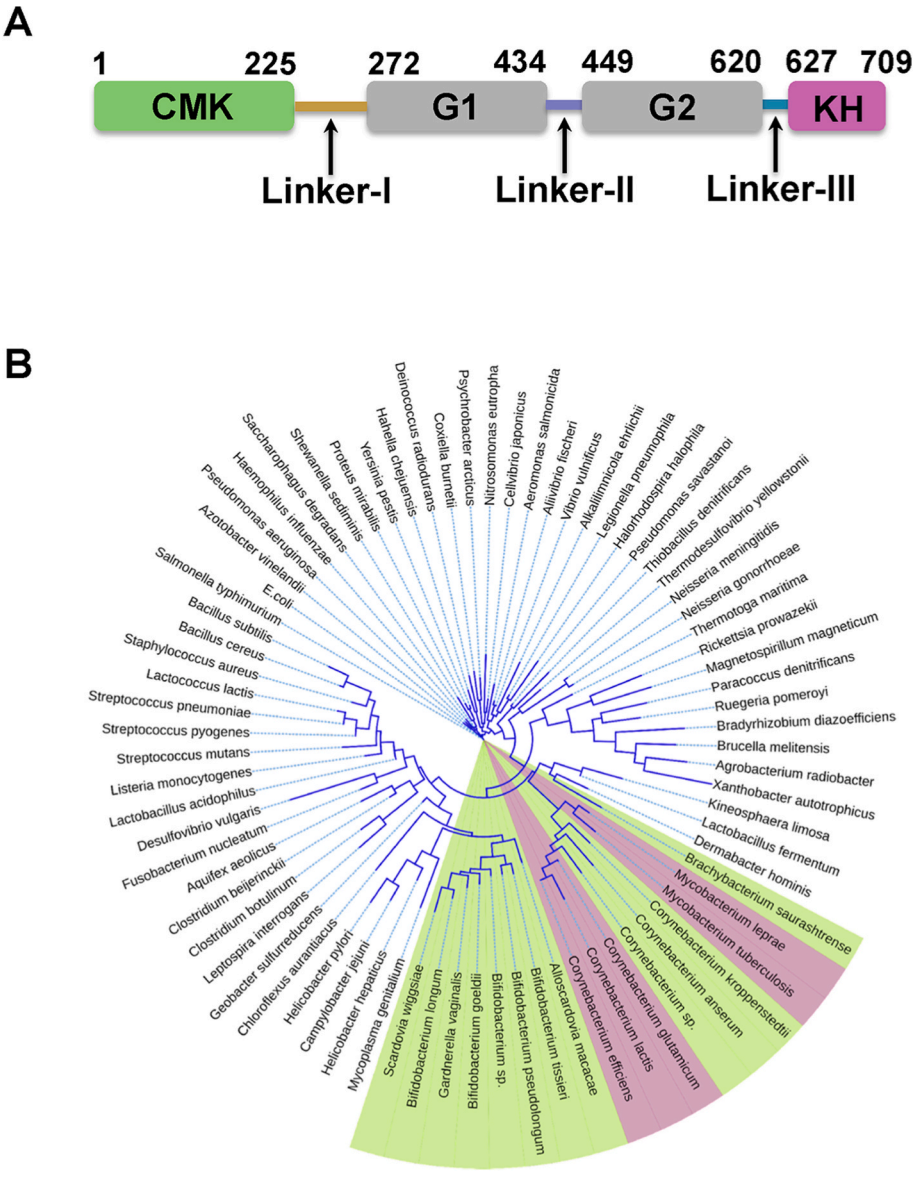


Fig. 1. Domain organization and phylogenetic relationship of BiCMKEngA. (A) Domains of BiCMKEngA. CMK domain (green), the linker region (yellow), GD1 (grey), GD2 (grey) and KH-Like domain (magenta) and the linker domain (blue). (B) Phylogenetic tree for EngA homologs from bacteria. The species highlighted in green show fused multi-domain (functional) fused EngA and the ones with unfused EngAs are highlighted in pink. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

BiCMK, respectively) of BiCMKEngA separately. However, we found that the constructs that exclude the linker region were insoluble (Supplementary Fig. S4) and hence both BiEngA and BiCMK contain the linker L1 (Supplementary Fig. S4). All three constructs were eluted as monomers and found to be stable in the pH range of 7–8. The MALS analysis further confirmed the molecular weight of the constructs to be 70kDa, 53 kDa 25kDa for BiCMKEngA, BiEngA, and BiCMK, respectively (Supplementary Figs. S5A, S5B, S5C). Since the stability and solubility of BiCMKEngA domains are largely influenced by the linker region, we surmised that this region might have a functional role in modulating the GTPase and Kinase activities of EngA and CMK domains, respectively.

Extensive efforts to crystallise BiCMKEngA did not yield success. We therefore resorted to homology modelling to gain structural insights (Fig. 2). CMK and EngA domains were modelled using the closest available homologous structures from *Mycobacterium smegmatis* (PDB ID; 3R20 and *Bacillus Subtilis* (PDB id: 4DCU), respectively. It is interesting to note that the 50 amino-acid linker region, connecting the CMK and EngA domains, could not be modelled with a higher confidence level due to the lack of homologous templates. However, *de novo* structure prediction methods, such as Robetta, modeled the linker as a short domain of two continuous helices, albeit with lower confidence (Fig. 2). This region seems to influence the overall stability and solubility of the protein as the domains lacking the linker were found to be insoluble (Supplementary Fig. S4).

3.3. The CMK domain of BiCMKEngA shows promiscuity for co-factors

Both full-length (BiCMKEngA) and the domain BiCMK exhibit kinase activity with similar catalytic efficiency for their cognate substrates, CMP and dCMP with the canonical phosphate donor ATP (Fig. 3A, Supplementary Table S3). Furthermore, BiCMK shows similar activity towards both CMP and dCMP, whereas *E.coli* CMK shows about 2-fold higher activity for dCMP than that for CMP [21]. Interestingly, like *E. coli* CMK, the BiCMK also utilises GTP as the phosphate donor. However, unlike *E.coli* CMK, the kinase activity of both CMP and dCMP reduces when GTP is used as the phosphate donor (Fig. 3A, Supplementary Table S3). Details of this data are shown in Supplementary Fig. S8.

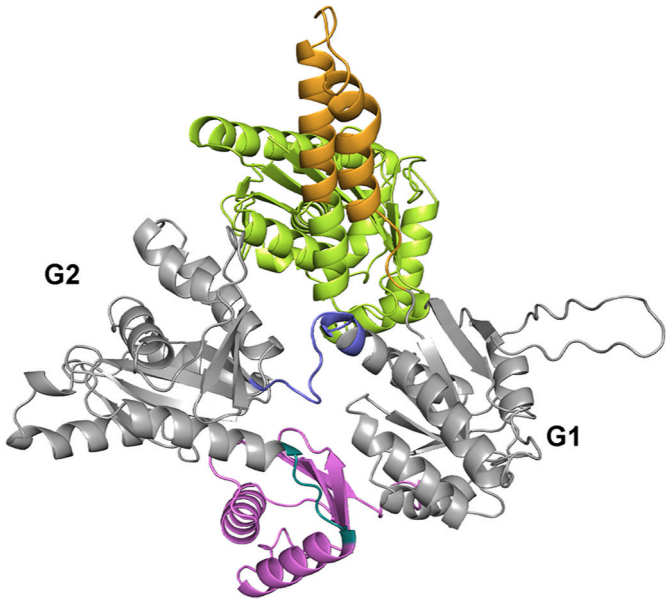


Fig. 2. Model structure of BiCMKEngA (Full-length protein). Different domains of the protein have been colored following Fig. 1A scheme; CMK domain (green), the linker region (yellow), GD1 (grey), GD2 (grey) and KH-Like domain (magenta) and the linker domain (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

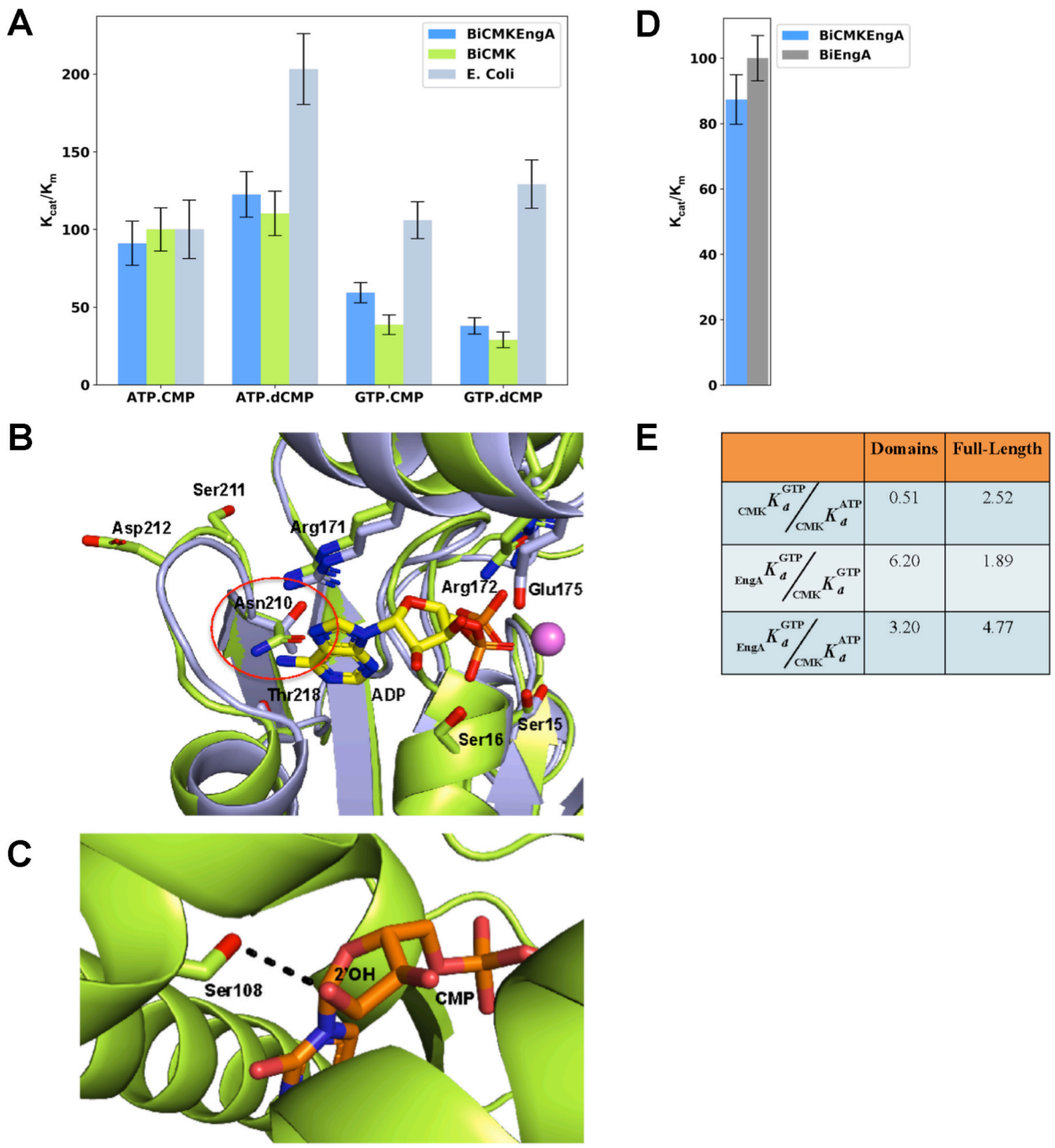
We rationalise the observation pertaining to the kinase activity of the CMK domain using homology modelling and docking studies. The ADP and CDP docked structure of BiCMKEngA shows that the residues corresponding to both the cofactor and the substrate binding sites are conserved (Fig. 3B). Furthermore, the catalytically important residues too, appear to be conserved. Our model shows that the amine group of the ATP base at the C^6 position lies in the vicinity of Asn_{210}, which can also stabilise the keto group of GTP present at the same position in the base (Fig. 3B). However, due to local electrostatic interactions, the longer side chain of Asn_{210} could be energetically unfavourable to accommodate GTP. Thus, making ATP a preferable phosphate donor than GTP for BiCMK. On the other hand, at the position equivalent to Asn_{210}, *E. coli* has a residue with a shorter side chain, Ser_{204}, which could accommodate both ATP and GTP equally efficiently. In other known bacterial CMKs, such as *Mycobacterium smegmatis*, this position is occupied by threonine (Thr_{190}), which seems to be incapable of accommodating the C^6 keto group of GTP, as this would require the residue to adopt an unfavourable conformation.

In BiCMK, in addition to the conserved interactions, the 2′OH of CMP interacts with Ser_{108}. This particular interaction is lost when CMP is replaced with dCMP, thus making the former a less preferred substrate for BiCMK (Fig. 3A and C and Supplementary Table S3). BiCMK does not show any detectable activity towards UMP, since the amine group of CMK base is stabilized by the conserved Ser/Thr (Thr_{32}) & Glu/Asp (Asp_{142}) and substitution of a keto group at this position (as seen in the case of UMP) would create unfavourable interaction (Supplementary Fig. S1).

3.4. Investigating mutual influence of BiCMK and EngA domains

Similar to the CMK domain, we measured the GTPase activity of the full-length BiCMKEngA and BiEngA (the EngA domain alone) and found them to be similar (Fig. 3D; Supplementary Fig. S8). Since the CMK domain also binds GTP (apart from ATP), from the ITC experiments (Supplementary Figs. S9–S12) it was expected that the full-length protein would show two GTP binding sites ; one at the CMK domain and another at GD1 of EngA in the presence of CMP. It is important to note that in all reported EngA purifications, GD2 was always pre-bound to GDP and its exchange requires regulatory factors or loosening of the active site with mild urea treatment followed by dialysis [33]. In line with this, we anticipated that in the ITC experiments using BiCMKEngA (Supplementary Fig. S11), we would observe GTP binding at two sites ; one at the CMK domain and another at the GD1 domain of EngA. However, these experiments indicate the existence of just one binding site, raising the possibility that the binding of GTP to the GD1 of the EngA domain is inhibited by the CMK domain and a regulatory event may be required for GTP to access this site. In contrast to this inference, the following data suggest that the GTPase activity by the EngA domain is not influenced by the CMK domain. The Kinetics experiments suggest that both BiEngA as well as BiCMKEngA exhibit similar GTPase activity (Fig. 3D, Supplementary Table S3). Also, the KD values for GTP, as determined from ITC experiments, were of the similar order (Supplementary Table S4).

Kinetics data shows that the CMK domain utilises both ATP and GTP as the phosphate donor (Supplementary Table S3). The Km, Vmax and Kcat for both BiCMKEngA and BiCMK are of similar orders, indicating that the presence of the EngA domain does not influence the activity of the CMK domain. The ITC thermograms with (and without) CMP suggest that CMP is required for the kinase domain to be functional (Supplementary Fig. S9, Fig. S10). Therefore, except for the one-site binding of GTP in the ITC experiments, the remaining data seems to indicate the absence of mutual influence of the CMK and EngA domains (of BiCMKEngA) on each other.



E

	Domains	Full-Length
$_{CMK}K_d^{GTP} / _{CMK}K_d^{ATP}$	0.51	2.52
$_{EngA}K_d^{GTP} / _{CMK}K_d^{GTP}$	6.20	1.89
$_{EngA}K_d^{GTP} / _{CMK}K_d^{ATP}$	3.20	4.77

Fig. 3. Catalytic activity of CMK domain (A) Kinase activities of various domains represented as a bar diagram. For details refer SI, Fig. S3 and Table S3 (B) Active site pocket around ATP and (C) CMP. Ligands are shown in sticks. Residues in the vicinity of C6 of ADP are highlighted with the circle. (D) GTPase activities of BiCMKEngA and BiEngA. (E) Relative dissociation constants of the full length and individual domains for ATP and GTP.

4. Discussion

Here we report a novel class of 'CMK fused EngA' protein with an evolutionary as well as functionally conserved GTPase EngA domain. The occurrence of CMKEngA is not well defined in the *Actinomycetes* phylum. The phytogenic proximity of both CMK and EngA genes (SI: Fig. S2) in the *Actinobacteria* shows the evolutionary gene fusion effects, which might have influenced these organisms to diverge as a commensal or pathogenic bacterium. EngA is known to exhibit nucleotide-dependent conformational changes and their influence on ribosome biogenesis [2]. The fusion of a kinase domain might be introducing further regulatory steps in ribosome binding and its maturation, which ultimately could influence the slow growth rate of *Bifidobacterium*. Continued studies in this direction will bring out the physiological importance of such a fusion event in future.

In this study, we attempted to explore the mutual regulation of different domains, CMK and EngA, of BiCMKEngA and observed that resolving this requires further investigations. An important finding that emerges from this work is that the CMK domain is not specific to ATP but can also use GTP as the phosphate donor to catalyse the conversion of CMP or dCMP to their respective diphosphates. However, the kinase activity is observed to be lower, when GTP is used as the phosphate donor. Usually, in a bacterial cell, the concentrations of ATP are higher (but of the same order) as that of GTP (3mM and 0.5mM, respectively) [34]. GTP is used to modulate specific metabolic processes like protein synthesis, signal transduction etc. Lower or limited use of GTP by kinases could therefore be a mechanism by which the depletion of the GTP pool is prevented. In some Kinases like the Adenylate Kinase (AdK) GTP binding to the active site leaves AdK in an unproductive state [35]. However, in BiCMK, GTP appears to bind in a productive conformation

and catalyse the reaction, albeit very slowly. This tempts us to speculate that basal CMK activity is important for this species, which may be achieved via GTP when the ATP pools are low.

Comparative analysis of bacterial CMK sequences, particularly with the well-characterised CMK from *E. coli*, provides insights on factors governing the promiscuity towards the phosphate donor. In general, the unfavourable interactions of the O6 carbonyl oxygen of the nucleobase with the backbone carboxylic groups of the loop at the base binding pocket were attributed as a major factor in discriminating between GTP and ATP by the kinases. In a few instances, additional interactions formed either by the side chains of the residues from the nucleobase binding pocket or a water bridge would override the unfavourable interactions posed by the backbone atoms, resulting in the binding of GTP to the kinase. In BiCMK, Asn 210 appears to make additional interactions to facilitate the binding of GTP.

An analysis of the available three-dimensional structures of EngA reveals that the enzyme primarily exists in two different conformations, which for the sake of convenience may be referred to as conformations A and B (Supplementary Figs. S13A and S13B); while A is represented by the structure of EngA from *B. subtilis* bound to GDP (PDB ID 4CDU), B is represented by the structure of EngA from *Coxiella burnetii* bound to GDP (PDB ID 5DN8). The difference between these is the position of the G-domains with respect to each other. It was also suggested that conformational changes such as these are driven by the nucleotides (GTP or GDP) occupied at GD1 and GD2. Supplementary Fig. S13 shows the model of BiCMKEngA superimposed on the two conformations A and B. Evidently, GD2 of conformation B clashes with the CMK domain of BiCMKEngA, indicating that if a similar conformational change (perhaps triggered by GTP hydrolysis at GD2) were to occur, the CMK domain must undergo a large conformational change; such a change will naturally be driven by the linker region if it forms a well-folded domain as shown in Fig. 2. All these indeed suggest an influence of the EngA domain on the CMK domain or vice-versa, and future experiments must be designed to demonstrate these.

EngA is found to be an essential protein for the survival of bacteria and its role in the maturation of ribosomes is well characterised. Structural and biochemical studies on *E.coli* EngA have shown that the protein exists in an autoinhibited state, which gets released upon hydrolysis of GTP, bound to the second GTPase domain, GD2. Perhaps, the GTP hydrolysis promotes the protein to adopt an extended conformation to facilitate ribosome binding [36]). However, in BiCMKEngA the role of the additional N –terminal CMK domain and the linker region in ribosome maturation is not clear. When BiCMKEngA was docked on *E.coli* EngA bound ribosome structure we could see that this CMK would introduce clashes with ribosomal proteins. Hence, we speculate that BiCMKEngA adopts a relatively larger extended conformation to interact with the ribosome as shown in Fig. 4. Furthermore, promiscuity of the CMK domain for the phosphate donor (GTP or ATP), hints at the utilization of GTP alone by all the nucleotide-binding domains of BiCMKEngA in the regulation of ribosome maturation. These aspects of BiCMKEngA merit further studies.

Bacterial EngAs have been considered potential drug targets against various diseases, including tuberculosis. *Bifidobacterium* is one of the early inhabitants of the human gut and an essential probiotic for the healthy host, CMK fusion to EngA might have intricate regulatory effects on ribosome biogenesis, which possibly could result in the slow growth rate of the organism as well as adaptation to the stringent gut conditions. The novel EngA from *Bifidobacterium*, reported here does open up new avenues in understanding species-specific EngAs and augment the development of pathogen-specific antibiotics that at the same time host protect microbiota.

5. Conclusion

BiCMKEngA is a unique variant of EngA GTPase exclusively found in a few *Actinomycetes* species. We were unable to demonstrate any inter-domain regulation of the protein using the kinetics studies of each domain, indicating further studies are required to understand the importance of CMK gene fusion in an essential protein like EngA.

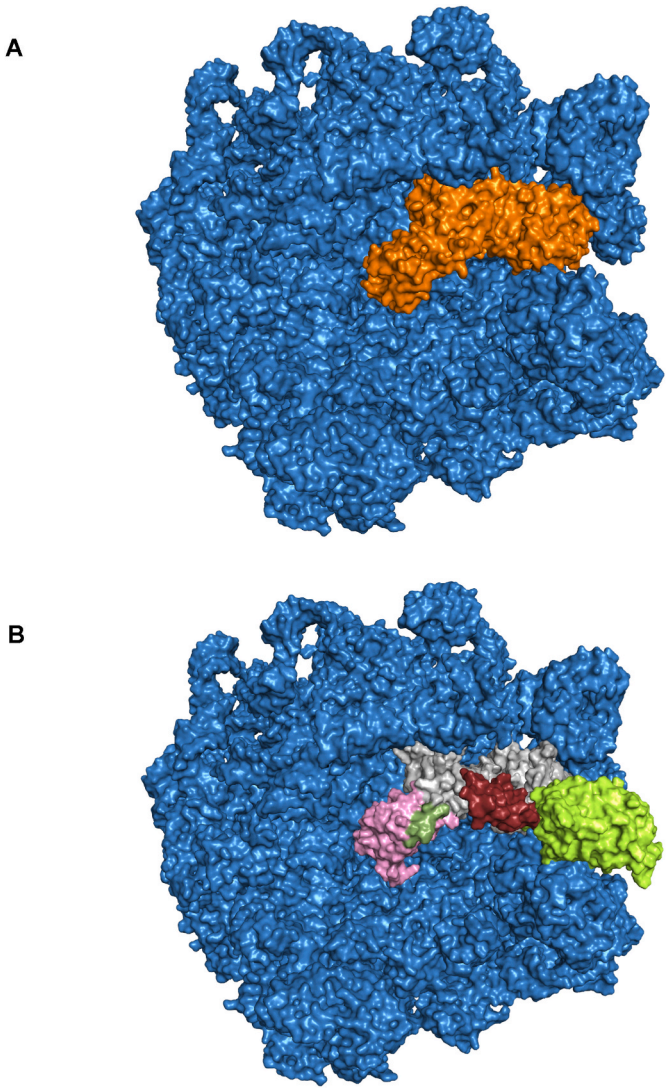


Fig. 4. Docking of BiCMKEngA on Ecoil EngA-ribosome structure. (A) Ribosome bound EngA structure from *E.Coli* (PDB:3J8G) (B) *E.Coli* Ribosome docked with BiCMKEngA model (3J8G is used as template).

Funding

This work was supported grant to KK and BP by the Department of Biotechnology, Government of India [grant number: BT/PR12502/BRB/10/1387/2015] and Funding from Ahmedabad University, via seed grant URBSASI21A3, respectively. AS acknowledges NET-CSIR for fellowship.

Author contribution

AS designed and performed the experiments; Study concept was conceived and designed by BP and KK; KK wrote the manuscript with inputs from AS and BP.

Declaration of competing interest

The authors declare no conflicts of interests.

Data availability

Data will be made available on request.

Acknowledgements

Authors thank Gayathri Pananghat (IISER- Pune), Dr. Subashchandrabose Chinnathambi, Dr.Moneesha Fernandes (CSIR-NCL) for providing excess to the ITC facility and Mr. Rajavee Srivastava for the last supplementary figure.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2022.101410.

References

[1] J. Hwang, M. Inouye, An essential GTPase, der, containing double GTP-bindin domains from *Escherichia coli* and *Thermotoga maritima*, J. Biol. Chem. 276 (33) (2001) 31415–31421, https://doi.org/10.1074/jbc.m104455200, 2001.

[2] X. Zhang, K. Yan, Y. Zhang, N. Li, C. Ma, Z. Li, Y. Zhang, B. Feng, J. Liu, Y. Sun, Y. Xu, J. Lei, N. Gao, Structural insights into the function of a unique tandem GTPase EngA in bacterial ribosome assembly, Nucleic Acids Res. 42 (21) (2014) 13430–13439, https://doi.org/10.1093/nar/gku1135, 2014.

[3] X. Ni, J.H. Davis, N. Jain, A. Razi, S. Benlekbir, A.G. McArthur, J.L. Rubinstein, R. A. Britton, J.R. Williamson, J. Ortega, YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit, Nucleic Acids Res. 44 (17) (2016) 8442–8455, https://doi.org/10.1093/nar/gkw678.

[4] A. Bharat, M. Jiang, S.M. Sullivan, J.R. Maddock, E.D. Brown, Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated Escherichia coli EngA, J. Bacteriol. 188 (22) (2006) 7992–7996, https://doi.org/10.1128/JB.00959-06.

[5] V.L. Robinson, J. Hwang, E. Fox, M. Inouye, A.M. Stock, Domain arrangement of Der, a switch protein containing two GTPase domains, Structure 10 (12) (2002) 1649–1658, https://doi.org/10.1016/s0969-2126(02)00905-x.

[6] S.P. Muench, L. Xu, S.E. Sedelnikova, D.W. Rice, The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding, Proc. Natl. Acad. Sci. USA 103 (33) (2006) 12359–12364, https://doi.org/10.1073/pnas.0602585103.

[7] H.K. Lamb, P. Thompson, C. Elliott, I.G. Charles, J. Richards, M. Lockyer, N. Watkins, C. Nichols, D.K. Stammers, C.R. Bagshaw, Functional analysis of the GTPases EngA and YhbZ encoded by *Salmonella typhimurium*, Protein Sci. 16 (11) (2007) 2391–2402, https://doi.org/10.1110/ps.072900907.

[8] S.K. Tomar, P. Kumar, S. Majumdar, V. Bhaskar, P. Dutta, B. Prakash, Extended C-terminus and length of the linker connecting the G-domains are species-specific variations in the EngA family of GTPases, FEBS Open bio 2 (2012) 191–195, https://doi.org/10.1016/j.fob.2012.07.009.

[9] E.M. Brown, M. Sadarangani, B.B. Finlay, The role of the immune system in governing host-microbe interactions in the intestine, Nat. Immunol. 14 (7) (2013) 660–667, https://doi.org/10.1038/ni.2611.

[10] G.E. Kaiko, T.S. Stappenbeck, Host–microbe interactions shaping the gastrointestinal environment, Trends Immunol. 35 (11) (2014) 538–548, https://doi.org/10.1016/j.it.2014.08.002.

[11] J.Z. von Martels, M.S. Sadabad, A.R. Bourgonje, T. Blokzijl, G. Dijkstra, K.N. Faber, H.J. Harmsen, The role of gut microbiota in health and disease: in vitro modeling of host-microbe interactions at the aerobe-anaerobe interphase of the human gut, Anaerobe 44 (2017) 3–12, https://doi.org/10.1016/j.anaerobe.2017.01.001.

[12] N. Agarwal, M. Pareek, P. Thakur, V. Pathak, Functional characterization of EngAMS, a P-loop GTPase of *Mycobacterium smegmatis*, PLoS One 7 (4) (2012), e34571, https://doi.org/10.1371/journal.pone.0034571.

[13] C.F. Beck, J. Neuhard, E. Thomassen, J.L. Ingraham, E. Kleker, *Salmonella typhimurium* mutants defective in cytidine monophosphate kinase (cmk), J. Bacteriol. 120 (3) (1974) 1370–1379, https://doi.org/10.1128/2Fjb.120.3.1370-1379.1974, 1974.

[14] J. Fricke 1, J. Neuhard, R.A. Kelln, S. Pedersen, The cmk gene encoding cytidine monophosphate kinase is located in the rpsA operon and is required for normal replication rate in Escherichia coli, J. Bacteriol. 177 (3) (1995) 517–523, https://doi.org/10.1128/jb.177.3.517-523.1995 (1995) Feb.

[15] A. Sorokin, P. Serror, P. Pujic, V. Azevedo, S.D. Ehrlich, The Bacillus subtilis chromosome region encoding homologues of the Escherichia coli mssA and rpsA gene products, Microbiology 141 (2) (1995) 311–319, https://doi.org/10.1099/13500872-141-2-311.

[16] L. Yu, J. Mack, P.J. Hajduk, S.J. Kakavas, A.Y. Saiki, C.G. Lerner, E.T. Olejniczak, Solution structure and function of an essential CMP kinase of Streptococcus pneumoniae, Protein Sci. 12 (11) (2003) 2613–2621, https://doi.org/10.1110/ps.03256803.

[17] D. Segura-Peña, N. Sekulic, S. Ort, M. Konrad, A. Lavie, Substrate-induced conformational changes in human UMP/CMP kinase, J. Biol. Chem. 279 (32) (2004) 33882–33889, https://doi.org/10.1074/jbc.m401989200.

[18] F. Fassy, O. Krebs, M. Lowinski, P. Ferrari, J. Winter, V. Collard-Dutilleul, K. Salahbey hocini, UMP kinase from Streptococcus pneumoniae: evidence for co-operative ATP binding and allosteric regulation, Biochem. J. 384 (3) (2004) 619–627, https://doi.org/10.1042/bj20040440.

[19] D.D. Leipe, E.V. Koonin, L. Aravind, Evolution and classification of P-loop kinases and related proteins, J. Mol. Biol. 333 (4) (2003) 781–815, https://doi.org/10.1016/j.jmb.2003.08.040.

[20] P. Briozzo, B. Golinelli-Pimpaneau, A.-M. Gilles, J.-F. Gaucher, S. Burlacu-Miron, H. Sakamoto, J. Janin, O. Bârzu, Structures of Escherichia coli CMP kinase alone and in complex with CDP: a new fold of the nucleoside monophosphate binding domain and insights into cytosine nucleotide specificity, Structure 6 (12) (1998) 1517–1527, https://doi.org/10.1016/s0969-2126(98)00150-6.

[21] C.P. Schultz, L. Ylisastigui-Pons, L. Serina, H. Sakamoto, H.H. Mantsch, J. Neuhard, O. Bârzu, A.M. Gilles, Structural and catalytic properties of CMP kinase from Bacillus subtilis: a comparative analysis with the homologous enzyme from Escherichia coli, Arch. Biochem. Biophys. 340 (1) (1997) 144–153, https://doi.org/10.1006/abbi.1997.9888.

[22] R. Mega, N. Nakagawa, S. Kuramitsu, R. Masui, The crystal structures of Thermus thermophilus CMP kinase complexed with a phosphoryl group acceptor and donor, PLoS One 15 (5) (2020), e0233689, https://doi.org/10.1371/journal.pone.0233689.

[23] The UniProt Consortium, UniProt: the universal protein knowledgebase in 2021, Nucleic Acids Res. 49 (D1) (2021) D480–D489, https://doi.org/10.1093/nar/gkaa1100.

[24] W. Li 1, L. Jaroszewski, A. Godzik, Clustering of highly homologous sequences to reduce the size of large protein databases, Bioinformatics 17 (3) (2001) 282–283, https://doi.org/10.1093/bioinformatics/17.3.282.

[25] Robert C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (5) (2004) 1792–1797, https://doi.org/10.1093/nar/gkh340, 2004.

[26] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2 - a multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (9) (2009) 1189–1191, https://doi.org/10.1093/bioinformatics/btp033.

[27] Glen Stecher, Koichiro Tamura, Sudhir Kumar, Molecular evolutionary genetics analysis (MEGA) for macOS, Mol. Biol. Evol. 37 (4) (2020) 1237–1239, https://doi.org/10.1093/molbev/msz312.

[28] Ivica Letunic, Bork Peer, Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, Nucleic Acids Res. 49 (W1) (2021) W293–W296, https://doi.org/10.1093/nar/gkab301.

[29] Jacques Oberto, SyntTax: a web server linking synteny to prokaryotic taxonomy, BMC Bioinf. 14 (4) (2013), https://doi.org/10.1186/1471-2105-14-4.

[30] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G.R. Lee, J. Wang, Q. Cong, L.N. Kinch, R.D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J.H. Pereira, A.V. Rodrigues, A.A. van Dijk, A. C. Ebrecht, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network, Science 373 (6557) (2021) 871–876, https://doi.org/10.1126/science.abj8754.

[31] C. Blondin, L. Serina, L. Wiesmuller, A.M. Gilles, O. Barzu, Improved spectrophotometric assay of nucleoside monophosphate kinase activity using the pyruvate kinase/lactate dehydrogenase coupling system, Anal. Biochem. 220 (1) (1994) 219–221, https://doi.org/10.1006/abio.1994.1326.

[32] A.D. Farr, P. Remigi, P.B. Rainey, Adaptive evolution by spontaneous domain fusion and protein relocalization, Nature ecology & evolution 1 (10) (2017) 1562–1568, https://doi.org/10.1038/s41559-017-0283-7.

[33] S.K. Tomar, N. Dhimole, M. Chatterjee, B. Prakash, Distinct GDP/GTP bound states of the tandem G-domains of EngA regulate ribosome binding, Nucleic Acids Res. 37 (7) (2009) 2359–2370, https://doi.org/10.1093/nar/gkp107.

[34] T.W. Traut, Physiological concentrations of purines and pyrimidines, Mol. Cell. Biochem. 140 (1) (1994) 1–22, https://doi.org/10.1007/bf00928361.

[35] P. Rogne, M. Rosselin, C. Grundström, C. Hedberg, U.H. Sauer, M. Wolf-Watz, Molecular mechanism of ATP versus GTP selectivity of adenylate kinase, Proc. Natl. Acad. Sci. USA 115 (12) (2018) 3012–3017, https://doi.org/10.1073/pnas.1721508115.

[36] C. da Silveira Tomé, A.E. Foucher, J.M. Jault, D. Housset, High concentrations of GTP induce conformational changes in the essential bacterial GTPase EngA and enhance its binding to the ribosome, FEBS J. 285 (1) (2018) 160–177, https://doi.org/10.1111/febs.14333.