Exploring the binding sites of *Staphylococcus aureus* phenylalanine tRNA synthetase: a homology model approach

Samar S. Elbaramawi[1, 2], Samy M. Ibrahim[2], El-Sayed M. Lashine[2], Mohamed E. El-Sadek[2], Efi Mantzourani[1], Claire Simons[1*]

[1]School of Pharmacy and Pharmaceutical Sciences, Cardiff University, Cardiff, CF10 3NB, United Kingdom. [2]Department of Medicinal Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, P.C. 44519, Egypt.

ABSTRACT:

Increased resistance of MRSA (multidrug resistance *Staphylococcus aureus*) to anti-infective drugs is a threat to global health necessitating the development of anti-infectives with novel mechanisms of action. Phenylalanine tRNA synthetase (PheRS) is a unique enzyme of the aminoacyl-tRNA synthetases (aaRSs), which are essential enzymes for protein biosynthesis. PheRS is an $(\alpha\beta)_2$ tetrameric enzyme composed of two alpha subunits (PheS) and two larger beta subunits (PheT). Our potential target in the drug development for the treatment of MRSA infections is the phenylalanine tRNA synthetase alpha subunit that contains the binding site for the natural substrate. There is no crystal structure available for *S. aureus* PheRS, therefore comparative structure modeling is required to establish a putative 3D structure for the required enzyme enabling development of new inhibitors with greater selectivity. The *S. aureus* PheRS alpha subunit homology model was constructed using Molecular Operating Environment (MOE) software. *Staphylococcus haemolyticus* PheRS was the main template while *Thermus*

thermophilus PheRS was utilised to predict the enzyme binding with $tRNA^{phe}$. The model has been evaluated and compared with the main template through Ramachandran plots, Verify 3D and Protein Statistical Analysis (ProSA). The query protein active site was predicted from its sequence using a conservation analysis tool. Docking suitable ligands using MOE into the constructed model was used to assess the predicted active sites. The docked ligands involved the PheRS natural substrate (phenylalanine), phenylalanyl-adenylate and several described *S. aureus* PheRS inhibitors.

Keywords:

MRSA; aminoacyl tRNA synthetase; computational analysis; molecular docking; binding site interactions.

INTRODUCTION

The aminoacyl tRNA synthetases (aaRSs) are essential enzymes for protein biosynthesis that catalyse the attachment of an amino acid to its cognate tRNA molecule in a two-step reaction. Cognate amino acids first react with ATP forming aminoacyl-adenylate, these activated forms of the amino acids are subsequently attached to their cognate tRNA by esterification. The catalytic steps are based on the ability of aaRSs to recognise amino acids, ATP and cognate tRNA for proper transformation of genetic information into proteins (1). Inhibition of either one of these stages results in accumulation of uncharged tRNA molecules, which bind to ribosomes causing interruption of polypeptide chain elongation and ultimately inhibition of protein synthesis (2). The aaRSs are categorised into two classes according to the structural features of the enzymes. Class I enzymes contain a Rossman fold in the catalytic core and two conserved

motifs, called HIGH and KMSKS. Class II enzymes have an antiparallel β-sheet with three conserved motifs in the catalytic centre (1, 3).

Increased resistance to anti-infective drugs is a threat to global health; hence anti-infective agents with novel mechanisms must be developed. *S. aureus* was chosen as MRSA is a growing problem in shared facilities such as hospitals, healthcare facilities and nursing homes. Studies indicate that the incidence of MRSA in the past few years have extensively increased worldwide. However, there are considerable differences between various countries. Whereas in the USA, Japan and southern European countries a high prevalence of MRSA between 20 and 60% exists, the prevalence in the Netherlands and Scandinavian countries is less than three percent (4).

AaRS provide a valuable target in bacterial protein biosynthesis, to date only one drug, mupirocin, which inhibits a specific type of aaRS (IleRS), has been licensed as a topical antibiotic for the treatment of methicillin-resistant *Staphylococcus aureus* (MRSA) (5). Since each amino acid requires an aaRS there is a potential wealth of targets suitable for rational drug design. The research described here focuses on PheRS. This target was selected because it is structurally unique among the aaRS enzymes and considerably different from human cytosolic and human mitochondrial aaRS, with 29.5% homology for PheRSα and 20.6% homology for PheRSβ (Supplementary Table S1). Phenylalanine tRNA synthetase (PheRS) is an $(\alpha\beta)_2$ tetrameric enzyme composed of two alpha subunits (PheS) and two larger beta subunits (PheT). According to the structure, PheRS is classified as a class II aaRSs as its catalytic domain is built around antiparallel β-sheets but functionally it resembles class I because it aminoacylates the 2' OH of the terminal ribose of tRNA whereas class II aminoacylate the 3' OH (6, 7).

There is no crystal structure available for *S. aureus* PheRS, therefore a homology model was constructed and subsequently structure and active site architecture analysed.

METHODS

Construction of the *S.aureus* PheS model

Homology search

The protein sequence for the alpha subunit (PheS) of *S. aureus* PheRS was obtained from the ExPASy proteomics server at the Swiss Bioinformatics Institute (8). The enzyme sequence for the alpha subunit has the Uniprot identifier P68849 (SYFA_STAAU) and is composed of 352 residues (9). Only PheS was used for the homology model construction as it contains the catalytic domain of the enzyme.

A homology search was performed using the SIB BLAST service (10, 11) accessible from the ExPASy server, which was used to align the query sequence (PheS) against the sequences in the protein data bank (12) and thus the close homologous proteins were identified. The alignment parameters and the thresholds, which were used for screening expected homologues, were used with their default values and BLOSUM62 comparison matrix. The phylogeny server (13) was used to build a phylogenetic tree for these homologous proteins, the query sequence and other PheS enzymes selected from different organisms.

Multiple sequence and structure alignment

The sequence of the query enzyme was aligned with the protein sequences of the most related PheRS templates: *Staphylococcus haemolyticus* (pdb: 2RHQ), *Pseudomonas aeruginosa (pdb:* 4P71), *Escherichia coli* (pdb: 3PCO) and *Thermus thermophilus* (pdb: 1PYS), using Clustal Omega 1.2.2 (14), as the local alignment of these sequences would reveal the most conserved

residues and the common features and motifs: motif 1, motif 2 and motif 3. This could be used to understand some of the expected structural and functional similarity between these enzymes. The secondary structure of *S. aureus* and the closest template (2RHQ) were determined using PSIPRED v3.3 (15).

3D model building

The molecular experiments were performed using Molecular Operating Environment (MOE) 2014.0901 molecular modeling software (16). Homology models were built using MOE-Homology using AMBER99 forcefield (17), which uses a dictionary to set the partial charges of atoms in amino acids. The final homology model was constructed using *S. haemolyticus* PheS (2RHQ) crystal structure. Ten intermediate models were generated and the final model was taken as the Cartesian average of all the intermediate models. All minimisations were performed until RMSD gradient of 0.05 kcal $mol^{-1}Å^{-1}$ with the specified forcefield and partial charges automatically calculated.

Model validation

Stereochemical quality of the polypeptide backbone and side chains was evaluated using Ramachandran plots obtained from the RAMPAGE server (18). The compatibility of the 3D model with its own amino acid 1D sequence was examined using Verify 3D (19). The ProSA server (20) was used to check defaults in the three dimensional protein structure based on statistical analysis. Validation data from the template (2RHQ) was used as the baseline to evaluate the model.

Docking

Phenylalanine as the natural substrate, phenylalanyl-adenylate (21) and known inhibitors with reported microbiological activity were built as ligands using MOE-Builder (22, 23), and the

energy was minimised for each ligand. These ligands were docked by running MOE-Dock with default setting: Placement: Triangular Matcher, Rescoring 1: London ΔG, 30 poses were constructed for each compound and the best scoring model-ligand complexes were selected. The ligand interactions within the constructed model were visualised using the MOE ligand interaction simulation.

Results and discussion

Homology model and validation

Initial screening for possible templates was performed using a BLAST analysis (10) of the amino acid sequence of *S. aureus* PheS, obtained from the ExPASy proteomics server (8), against the PDB resolved structures, four structures were identified for consideration as possible templates (Table 1).

Table 1 The first four hits in the *S. aureus* PheRS-α subunit BLAST results

Organism	**PDB code**	**BLAST[a] score**	**Sequence identity[b]**	**Sequence identity%**	**Positive %**	**Chain length**	**E-Value**
Staphylococcus haemolyticus	2RHQ	536	250/268	93	98	294	0.0
Pseudomonas aeruginosa	4P71	372	177/340	52	69	338	$5e^{-127}$
Escherichia coli	3PCO	341	169/338	50	67	327	$5e^{-115}$
Thermus thermophilus	1PYS	270	147/359	41	60	350	$5e^{-87}$

[a] The BLAST score for an alignment is calculated by summing the scores for each aligned position and the scores for gaps. [b] (Number of identical residues)/(length of sequence fragment identified by PSI-BLAST).

For a structure to be considered a template, it should be wild-type, rather than mutant or engineered, have a reasonable sequence identity with *S. aureus* PheRS and also have the same function. The first four native hits are bacterial PheRS enzymes. The PheRS enzyme of *Staphylococcus haemolyticus* (24) was the best template due to very high sequence identity (93%).

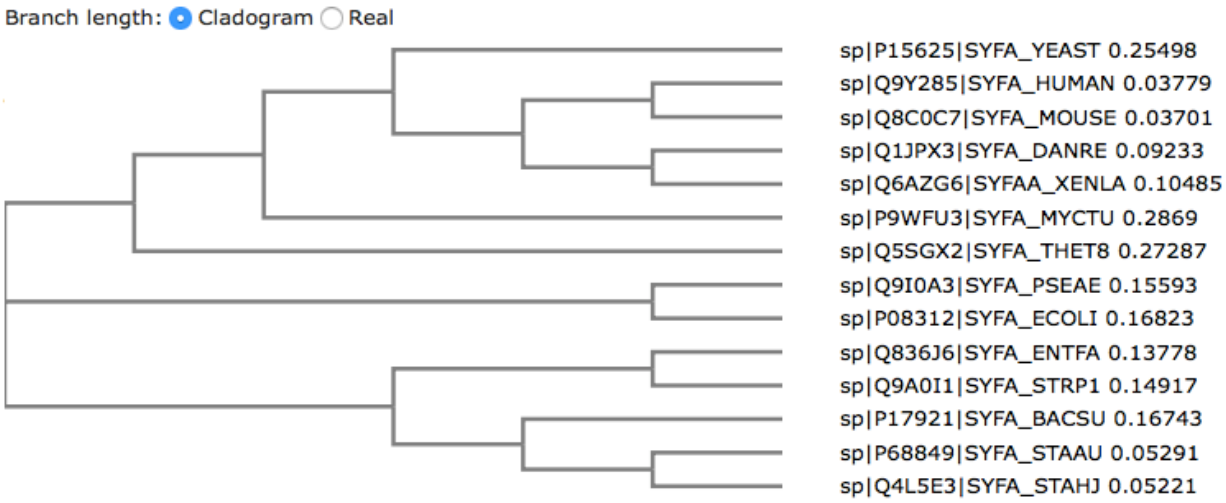


Figure 1 The phylogenetic tree of *S. aureus* PheS in relation to PheS enzyme from other organisms: *Saccharomyces cerevisiae* (yeast) (P15625); human (Q9Y285); mouse (Q8C0C7); zebra fish (Q1JPX3); African clawed frog (Q6AZG6); *Mycobacterium tuberculosis* (P9WFU3); *Thermus thermophilus (*Q5SGX2*); Pseudomonas aeruginosa* (Q9I0A3); *Escherichia coli* (P08312); *Enterococcus faecalis* (Q836J6); *Streptococcus pyogenes* (Q9A0I1); *Bacillus subtilis* (P17921); *Staphylococcus aureus* (P68849) and *Staphylococcus haemolyticus* (Q4L5E3*).*

To obtain more information related to the best possible template, the phylogeny server (13) was used to construct a phylogenetic tree using PheRS protein sequences from different organisms to determine the relative distances between the these enzyme and the query sequence (Fig. 1). The phylogenic tree demonstrates the different evolutionary branching of the prokaryotic and eukaryotic PheRS enzyme. The closest homologies to *S. aureus* in this group of species are: the Gram-positive bacteria *S. haemolyticus* (STAHJ) followed by *Bacillus subtillis* (BACSU), *Enterococcus feacalis* (ENTFA) and *Streptococcus pyogenes* (STRP1) respectively.

A lower homology is observed with the Gram-negative *E. coli* and with *Mycobacterium tuberculosis*, with clear difference observed for non-bacterial organisms in agreement with the phylogenic tree. Clustal analysis and percent identity (Supplementary Fig. S2 and S3) provide further validation.

Multiple Sequence and structural alignments

Clustal Omega 1.2.2 (14) was used to align the preferred template sequences and the query sequence of *S. aureus* PheS (Fig. 2). Conservation was clearly observed between most of the residues of the query sequence and the closely related template *S. haemolyticus*.

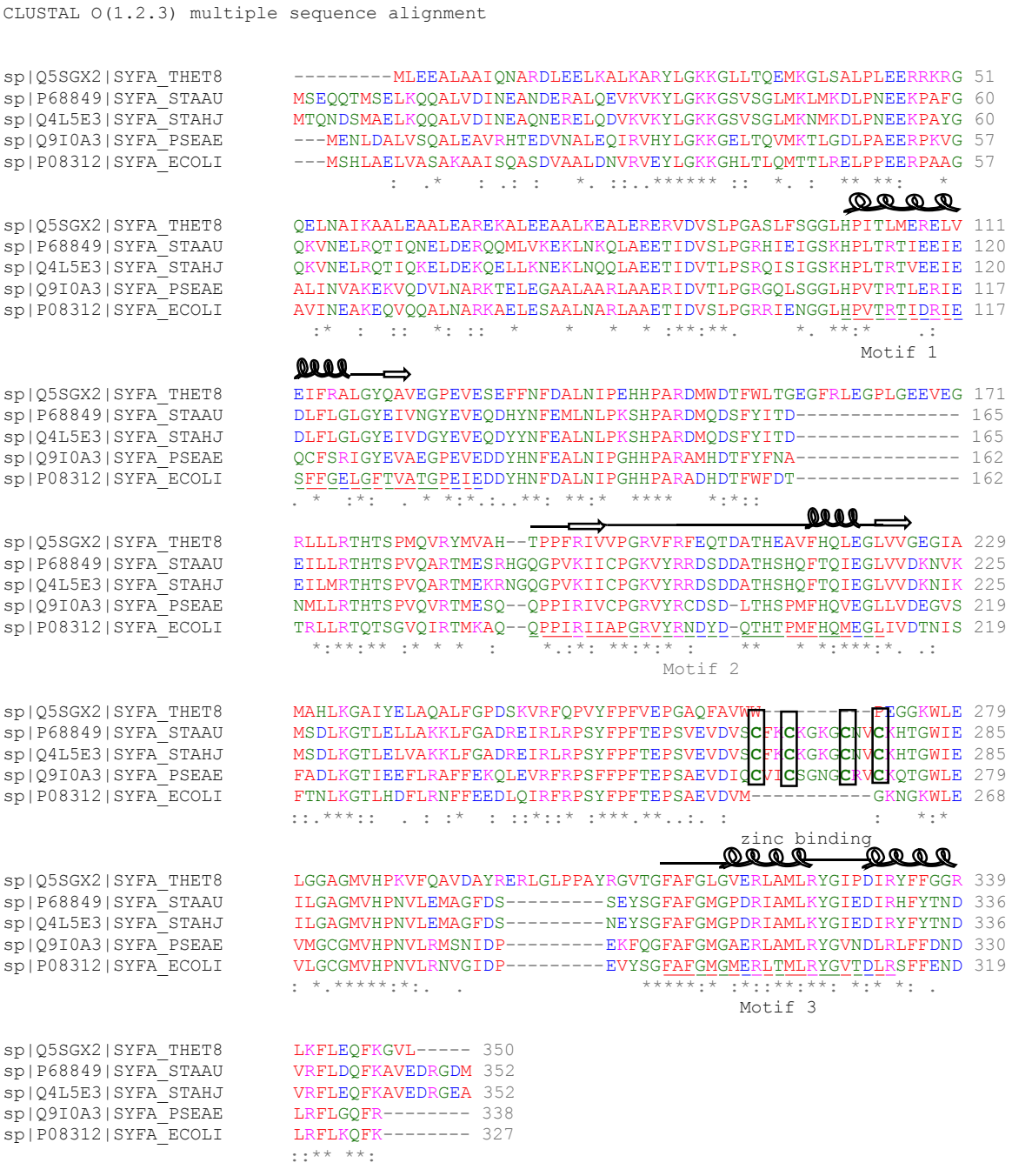


Figure 2 Sequence alignment of PheS enzymes of *Thermus thermophilus*, *Staphylococcus aureus*, *Staphylococcus haemolyticus*, *Pseudomonas aeruginosa* and *Escherichia Coli* using Clustal O in which "*" means that the residues are identical, ":" means that conserved substitutions have been observed, "." means that semi-conserved substitutions are observed. The residues are coloured according to their chemical properties where red, small hydrophobic (AVFPMILWY); blue, acidic (DE); purple, basic (RHK); green, hydroxyl + amine + basic (STYHCNGQ).

The PheRS α-subunit contains three conserved sequence motifs (motif 1, motif 2 and motif 3), which are typical for class II aminoacyl tRNA synthetases (9). These motifs play an important role in the aminoacylation process, as they are responsible for active site formation. In this active pocket, binding occurs with ATP/phenylalanine as well as interaction with the acceptor stem of $tRNA^{phe}$. All the residues in motifs 1, 2 and 3 that form the active site in both *S. aureus* and *S. haemolyticus* are conserved. PSIPRED (15) secondary structure prediction for *S. aureus* PheS showed that motif 1 (110-136) contained a high helix content surrounded by a short strand and coil, motif 2 (196-218) folds mainly in coils and short helix and strands and motif 3 (310-330) at the C terminal of the α-subunit clearly folds to coils and helices. The crystal structure of *S. haemolyticus* (2RHQ) shows these folds indicating good agreement with the predictions of PSIPRED for the query sequence (Supplementary data Fig. S4).

3D homology model and validation

The homology model was constructed using *S. haemolyticus* PheRS-α subunit (2RHQ) crystal structure through Molecular Operating Environment (MOE) software. The constructed model was subjected to a number of checks to assess its quality. Stereochemical quality was evaluated by Ramachandran plots using the RAMPAGE server (18) and the compatibility of the 3D model with its own amino acid 1D sequence was validated using Verify 3D (19). Overall protein structure was evaluated using ProSA (20).

Validation results would propose that the model performed well in terms of the backbone stereochemistry and amino acid environment. In the Ramachandran plot, a total of 99.7% of the residues were in the allowed region, which compared well with the template 2RHQ (98.9%),

revealing that the main chain dihedral ϕ and ψ angles in the model were reasonably accurate (Supplementary Fig. S1). Only one amino acid residue (Gly272) was found in the outlier region in the model, which is away from the active site and would not be expected to affect enzyme function (Fig. 9).

Verify 3D evaluated three parameters for each amino acid, which are secondary structure, degree of buried surface area and fraction of side chain area that is covered by polar atoms. These three parameters were evaluated for each residue in the structure and a correlation was calculated between this set of observed parameters and the ideal parameters of the amino acid type to which it has been assigned. Verify 3D should stay above 0.2 and not fall under zero (19). The percentage of residues, which are more than 0.2 was 85.25% and 87.77% for the model and the template, respectively. However, the percentage of residues that are less than zero was 9.71% and 7.04% for the model and the template, respectively. The model residues that fall under zero are the first 27 residues, which are away from the active site.

Protein statistical analysis (ProSA) provides two plots; the first (Fig. 3a) shows the local model quality by plotting energies as a function of amino acid sequence position, generally positive values relate to erroneous parts of the input structure. The second plot (Fig. 3b) indicates overall model quality from which the z-score is calculated, its value is displayed in a plot which has the z-scores of all experimentally determined protein chains in the current PDB determined by X-ray crystallography or NMR spectroscopy, a negative score shows a good model while a positive value would shows errors. The z-score of the model was -7.72 compared with the z-score of the template 2RHQ which was -7.73.

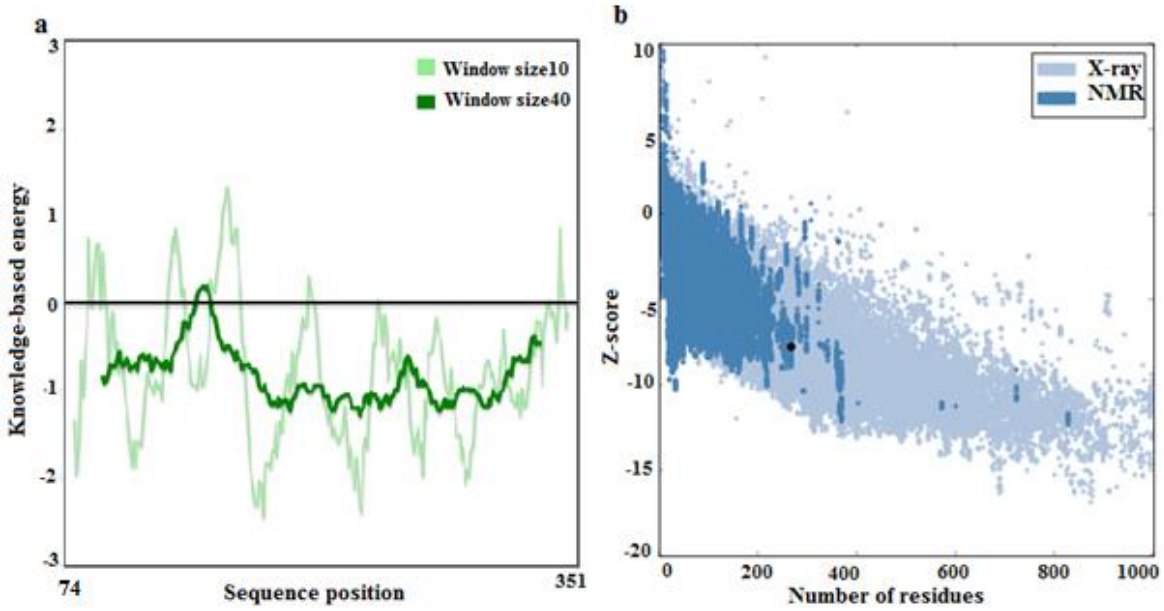


Figure 3 ProSA output for the *S. aureus* PheRS-α subunit model. a: shows the local model quality by plotting energies as a function of amino acid sequence position. b: shows the overall model quality by calculating z-score (dark spot).

Moreover, superimposition of the model with the main template (2RHQ) using MOE showed a low RMSD of 0.960 Å indicating a high degree of similarity.

Model validation using Ramachandran plot, Verify 3D and ProSA indicated a good model in terms of quality of backbone and side-chain stereochemistry for *S. aureus* PheS. Further validation of the active site architecture was performed by natural substrate and ligand docking experiments.

Active site validation and docking

Validation of the predicted active site was assessed through Clustal O multiple sequence alignment, the MOE alignment service and by docking known ligands into the constructed model. The active site is a large hydrophobic pocket formed by the following residues Leu148, His152, Ala154, Asp159, Ser160, Thr171, His172, Ser174, Gln177, Met181, Arg200, Phe212,

Gln214, Glu216, Leu218, Tyr253, Phe254, Pro255, Phe256, Thr257, Glu285, Leu287, Ile286, Gly288, Ala289, Gly290, Val292, Val296, Ala311, Phe312, Gly313 and Arg318. Characteristic with other bacterial enzymes, one wall of the phenylalanine binding pocket contains amino acids that can participate in hydrogen bonding and electrostatic interaction while the opposite was contains mainly hydrophobic amino acids (25). Through docking of the natural ligand, phenylalanine, hydrogen bonding interaction of the phenylalanine positively charged amine group and the acidic amino acid Glu216 was observed (Table 2) on one face of the pocket, while recognition on the hydrophobic side of the active site is via two phenylalanine residues, Phe254 and Phe256, from the eubacteria-specific FPF loop (25).

In contrast the FPF triad is absent from the phenylalanine active site in all eukaryotic PheRS, in human PheRS a triplet of residues Asn410, Pro411 and Tyr412 replace the FPF loop (Supplementary Fig. S2) (26). This difference can be exploited in selective drug design.

Table 2 3D and 2D models of binding interactions of phenylalanine in *S. aureus* PheRS active site

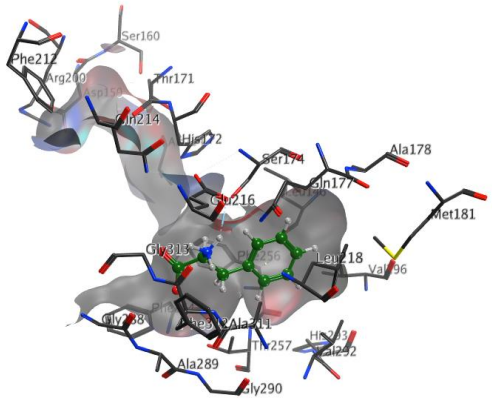
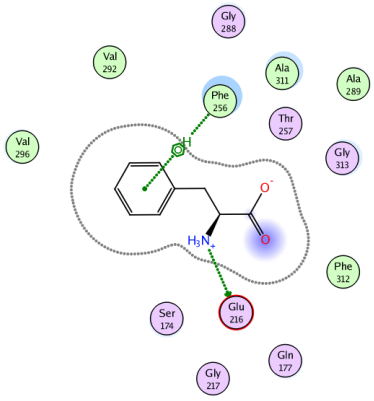
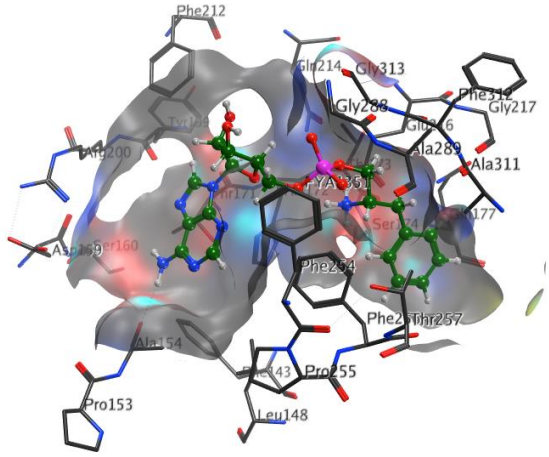
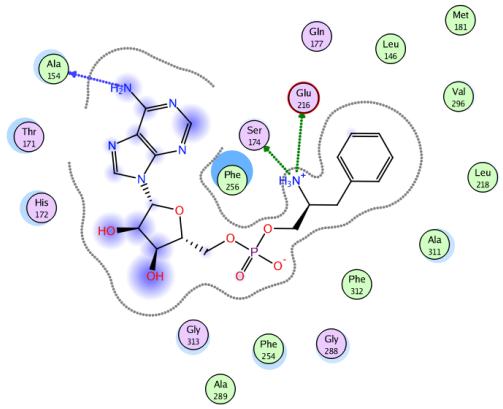
3D structure	2D ligand interactions
	

Table 3 3D and 2D models of binding interactions of phenylalanyl-adenylate in *S. aureus* PheRS active site

3D structure	2D ligand interactions
	

The same hydrogen bonding interaction between the amino group of phenylalanine and the acidic Glu216 was noted as a key binding interaction on docking phenylalanyl-adenylate (21) as well as hydrogen bonding interaction with Ser174, with the adenylate moiety positioned further along the hydrophobic pocket channel. The adenine ring may form π-stacking interactions with Phe212 and an additional H-boning interaction between the adenine amino group and Ala154 may help with correct orientation in the active site (Table 3).

Two classes of *S. aureus* PheRS inhibitors were investigated through docking studies, phenyl-thiazolylurea-sulfonamides (22) and ethanolamine derivatives (23). As observed for phenylalanine and phenylalanyl-adenylate, the acidic amino acid residue Glu216 forms H-bonding interaction with the ligands acting as an anchor and directing correct orientation within the hydrophobic pocket. Interaction between Glu216 and the phenyl-thiazolylurea-sulfonamides was though the urea NH moiety and/or the thiazoyl sulfur group (Table 4), while for the ethanolamine derivatives the interaction was observed via the ethanolamine OH or NH of the linker chain (Table 5).

Table 4 Representative 3D and 2D models of binding interactions of phenyl-thiazolylurea-sulfonamides in *S. aureus* PheRS active site.

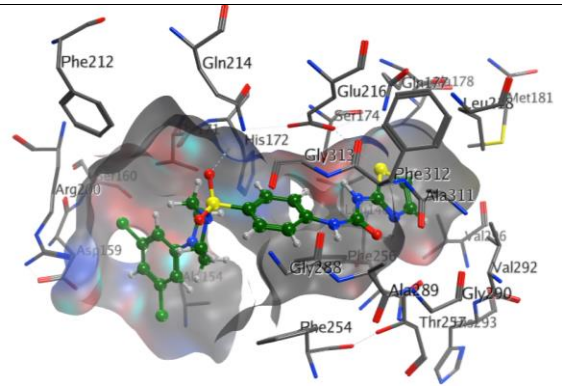
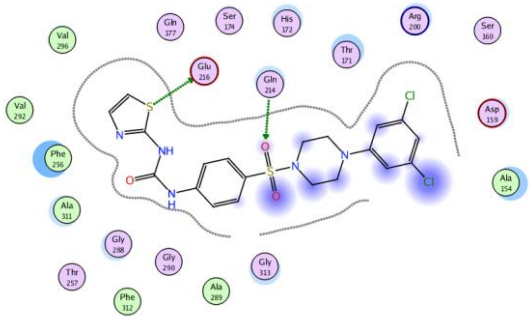
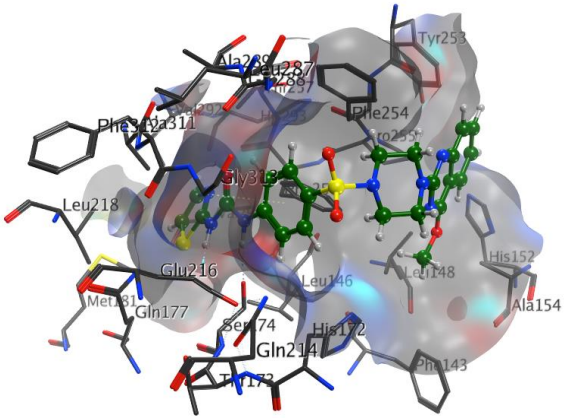
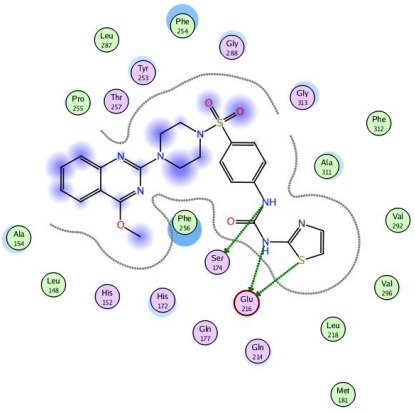
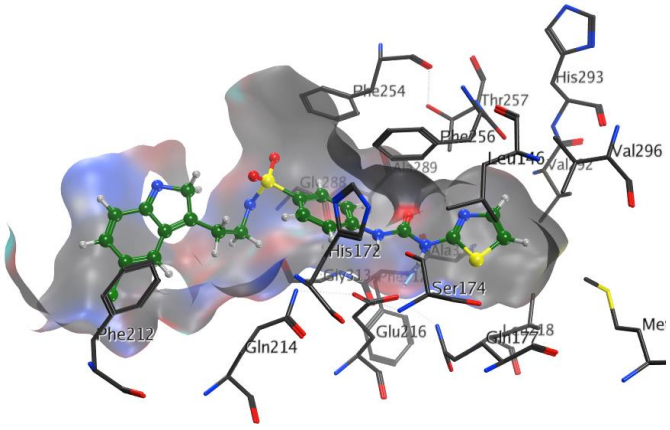
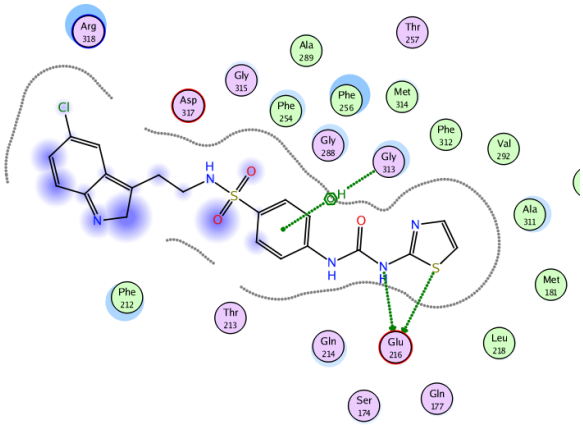
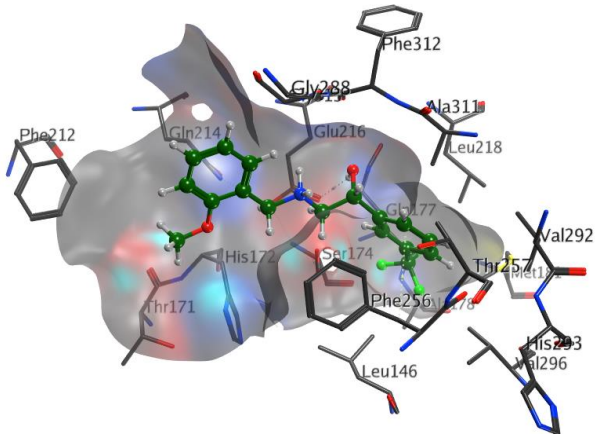
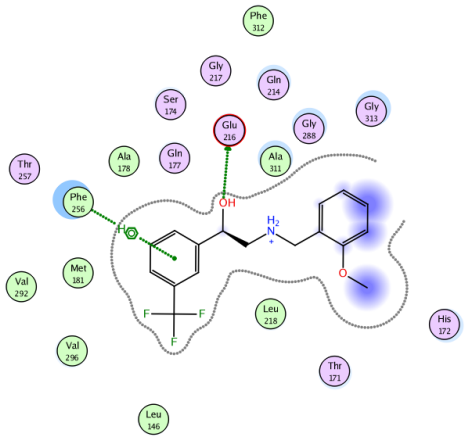
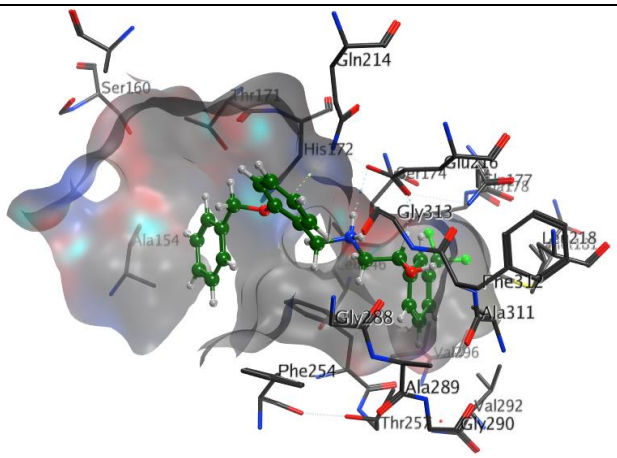
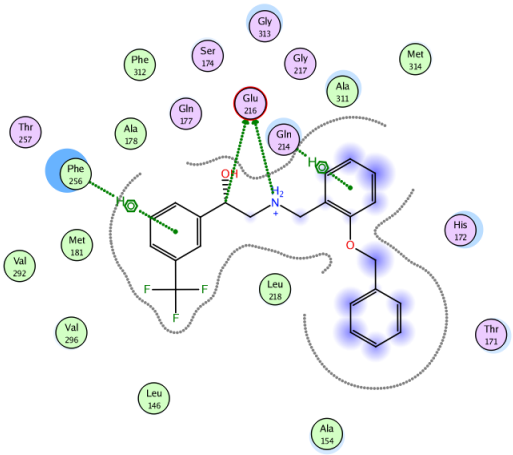
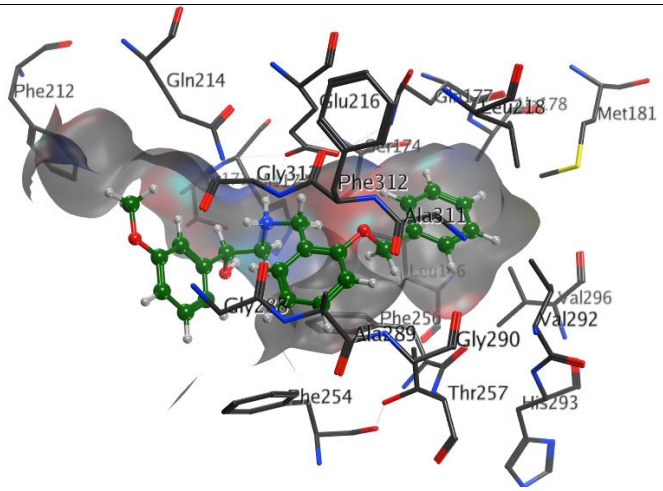
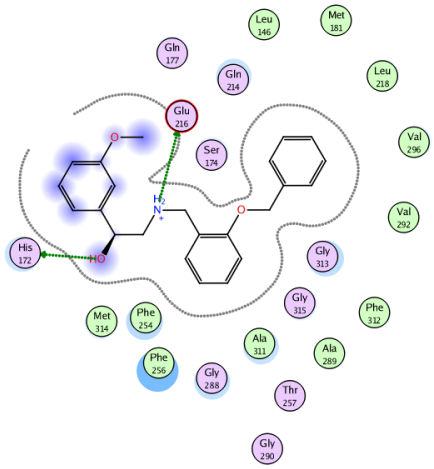
3D structure	2D ligand interactions
	
	
	

Table 5 Representative 3D and 2D models of binding interactions of ethanolamine inhibitors in *S. aureus* PheRS active site

3D structure	2D ligand interactions
	
	
	

Additional H-bonding interactions observed for some of the inhibitors and the enzyme active site included interactions with Ala154, His172, Ser174, Phe256, Glu285 and Gly288. The binding mode in the *S. aureus* PheS model was comparable with the binding of ligands in *S. haemolyticus* (2RHQ); indicating the conservation of active site key residues.

Metal Binding

Zinc metal

An experimental study on *S. haemolyticus* PheRS by Artem *et al.* (24) suggested that a zinc finger is present, with residues in the region of Ser266 to Gly282 forming a compact metal binding site where the four sulfur ligands related to Cys267, Cys270, Cys275 and Cys278 form a tetrahedral coordination with zinc (Fig. 4).

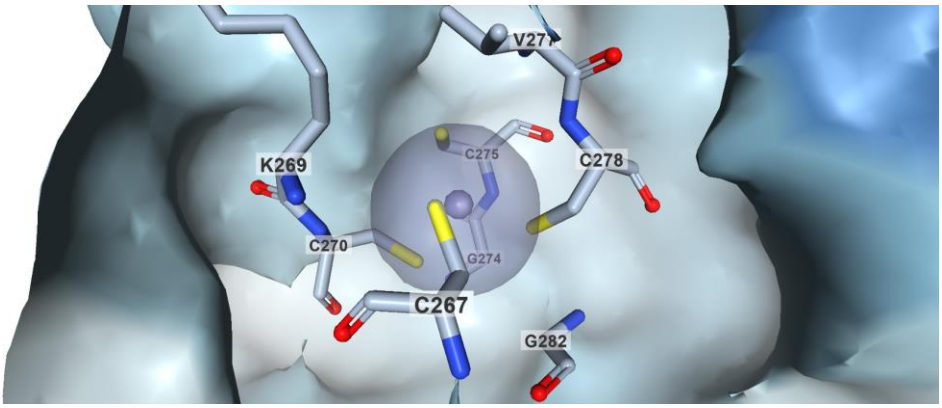


Figure 4 Zinc binding site in *S. haemolyticus* PheRS-α subunit (http://www.ebi.ac.uk/pdbe)

All PheRS-α subunit proteins from Gram-positive organisms have a conserved sequence insertion in the vicinity of residue 270. Within this insertion, there are the four cysteine residues which bind to zinc metal (Fig. 2). This was supported by the alignment of *S. aureus* PheRS α-subunit with the suggested template where all four conserved cysteine residues are identified in the region of this domain in the query sequence (Fig. 5). The zinc metal is not involved in the

active binding site and is also positioned far from the putative $tRNA^{phe}$ binding domains (Fig. 10).

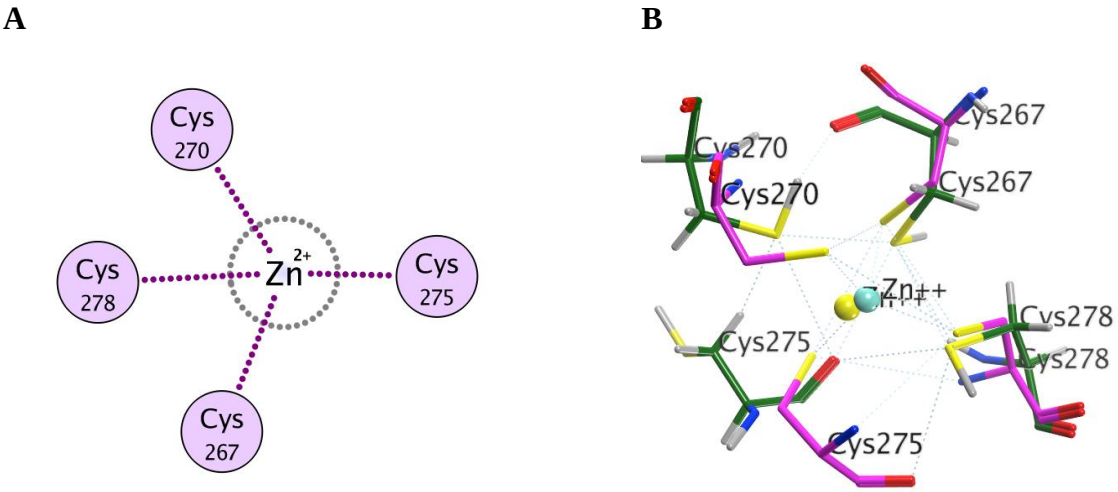


Figure 5 (**A**) Two dimensional form of tetrahedral coordination of zinc with four cysteine residues in the *S. aureus* PheRS-α subunit model. (**B**) Superimposition of *S. aureus* PheRs α-subunit model residues (green) and *S. haemolyticus* template (2RHQ, purple) residues interacting with the zinc ion (grey sphere *S. aureus* model, cyan sphere *S. haemolyticus*).

The metal binding site at the α/β- subunit interface

The magnesium ion is located at the interface between the α and β-subunits by five electrostatic interactions with acidic residues Asp β452, Glu β461 and Glu β462 from the beta subunit and with Glu α262 from the alpha subunit in proximity to the active site. Moreover, there is a sixth weak interaction of the metal ion with the carbonyl-oxygen atom of Asp β458 (27). Two magnesium ions are present per $(\alpha\beta)_2$ tetramer. The metal-mediated residues are conserved in *T. thermophilus*, *S. haemolyticus* and *S. aureus* (Table 6). Although the alpha subunit is responsible for the catalytic activity for binding with ATP/phenylalanine, biochemical studies (27) have shown that the alpha subunit is not able to catalyse the first step of the reaction in the absence of the beta subunit.

Table 6 Amino acids responsible for magnesium binding at the α/β- subunit interface

Thermus thermophilus	*Staphylococcus haemolyticus*	*Staphylococcus aureus*
Asp β452	Asp β 461	Asp β 461
Asp β 458 (via carbonyl-oxygen)	Asp β 467 (via carbonyl-oxygen)	Asp β 467 (via carbonyl-oxygen)
Glu β 461	Glu β 470	Glu β 470
Glu β 462	Glu β 471	Glu β 471
Glu α 262	Glu α 258	Glu α 258

Recognition and binding of PheRS with cognate tRNA^{phe}

The tRNA crystal structure binding with *S. haemolyticus* PheRS is not available therefore; tRNA^{phe} binding with *S. aureus* PheRS has been predicted from the published X-ray structure of *Thermus thermophilus* PheRS complexed with tRNA^{phe} (21, 28, 29). From the published crystal structure of *Thermus thermophilus*, it was found that the β-subunit contains eight complex domains. β1 (residues 1-39 and 152-183) and β5 (residues 401-473) domains have the same topology comprised of helix-loop-helix motifs which exhibit interaction with certain DNA sequences of *Thermus thermophilus* genomic DNA (30). However, the *S. aureus* PheRS amino acid sequence does not show this helix-loop-helix in the corresponding area so *S. aureus* enzyme is not predicted to bind with DNA (9). The β2 domain (residues 40-151) is structurally related to the anticodon binding domain of the bacterial aspartyl-tRNA synthetase (31) while the β8 domain (residues 692-785) resembles the RNA binding domain of spliceosomal protein U1A (32). Experimental data (26, 33) shows that the tRNA^{phe} anticodon tends towards β8 instead of

β2. The β6 domain (residues 482-514 and residues 564-679) together with the β7 domain (residues 515-563), whose function is still ambiguous, are considered as pseudocatalytic domains as they are structurally related to the active site of the α-subunit (29). The β3 (residues 211-264 and residues 329-396) and β4 (residues 265-328) domains resemble the Scr homology (SH3) domain, which occur in large numbers in eukaryotic tyrosine kinase but their function is still unknown (34). Overall, the C-terminal β8 domain is the essential one in the PheRS β-subunit for recognition and binding with cognate $tRNA^{phe}$ and other β-subunit domains have no clear role in the aminoacylation process.

The PheRS enzyme binds two $tRNA^{phe}$ molecules. Each tRNA binds across all four subunits of the enzyme, which indicates the necessity of the $\alpha_2\beta_2$ quaternary architecture for the enzyme activity (28). The CCA-end and the acceptor stem of the $tRNA^{phe}$ molecule interact with the α-subunit active site and with the N-terminal domain of the β-subunit for the same heterodimer. The anticodon loop of $tRNA^{phe}$ is recognised by the β8 domain, which is located at the C-terminal domain of the β*-subunit of the second heterodimer (* for the second heterodimer).

Table 7 Homology of *S. aureus and S. haemolyticus* and *T. thermophilus* PheRS, through SIB BLAST service

Organism	**PDB code**	**BLAST score**	**Sequence identity%**	**Positive %**	**Chain length**	**E-Value**
S. haemolyticus	2RHS	1258	78	90	800	0.0
T. thermophilus	1PYS	388	34	51	785	$2e^{-121}$

Table 7 shows homology between *S. aureus*, *S. haemolyticus* and *T. thermophilus*, through SIB BLAST service (10, 11) accessible from ExPASy server. *S. aureus* PheRS-β subunit (PheT) has the uniprot identifier Q9AGR3 (SYFB_STAAU) composed of 800 residues.

Proposed interactions of the GAA anticodon and CCA end in $tRNA^{phe}$ with *S. aureus* PheRS residues are based on the Clustal alignment of both α and β-subunits for the query enzyme with *S. haemolyticus* and *T. thermophilus* (Fig. 2, 6) where boxed amino acids residues at the C terminal β8 domain in Fig. 6 are responsible for recognition and interaction with $tRNA^{phe}$ anticodon bases and also their genomic alignment of $tRNA^{phe}$ (Fig. 7). The sequence conservation degree between *β*-subunits for *S. aureus* and *T. thermophilus* is 34% identity and 51% similarity. Although these percentages are lower than that between *S. aureus* and *S. haemolyticus*, the available *T. thermophilus* crystal structure was used to predict binding of tRNA with the enzyme, as importantly the predicted binding residues are conserved between PheRS of these three bacteria.

Figure 6. Sequence alignment of PheT (β-subunit) enzymes of *Staphylococcus aureus*, *Staphylococcus haemolyticus* and *Thermus thermophilus*, using Clustal Omega in which "*" means that the residues are identical, ":" means that conserved substitutions have been observed, "." means that semi-conserved substitutions are observed. The residues are coloured according to their chemical properties where red, small hydrophobic (AVFPMILWY); blue, acidic (DE); purple, basic (RHK); green, hydroxyl + amine + basic (STYHCNGQ). Boxed amino acid residues recognise the $tRNA^{phe}$ anticodon.

The tRNA^{phe} sequence of *S. aureus* is available from the National Centre for Biotechnology Information (www.ncbi.nlm.gov). The tRNA^{phe} sequence length for *S. aureus* and *S. haemolyticus* is 73 bases while *T. thermophilus* tRNA^{phe} has 76 bases. Clustal Omega 1.2.3 for RNA (13), is utilized in alignment with tRNA^{phe} for *S. aureus, S. haemolyticus and T. thermophilus* as the RNA alignment of these bases would reveal the most conserved nucleotides, specially for the anticodon (GAA) and CAA end playing a vital role in binding with PheRS (Fig. 7).

```
CLUSTAL O(1.2.3) multiple sequence alignment

gi|88193823:c1864867-1864795     GGTTCAGTAGCTCAGTTGGTAGAGCAATGGATTGAAGCTCCATGTGTCGGCAGTTCGACT
gi|70725001:1140750-1140822      GGTTCAGTAGCTCAGTTGGTAGAGCAATGGATTGAAGCTCCATGTGTCGGCAGTTCGACT
gi|55979969:1311554-1311629      GCCGAGGTAGCTCAGTTGGTAGAGCATGCGACTGAAAATCGCAGTGTCGGCGGTTCGATT
                                 *     ********************   ** ****  **   ******** ****** *

gi|88193823:c1864867-1864795     CTGTCCT---GAACAA
gi|70725001:1140750-1140822      CTGTCCT---GAACCA
gi|55979969:1311554-1311629      CCGCCCCTCGGCACCA
                                 * * **    * ****
```

Figure 7 Genomic alignment of tRNA^{phe} for *S. aureus, S. haemolyticus* and *T. thermophilus*, respectively, where boxed nucleotides are conserved GAA anticodon and CCA end.

Anticodon (GAA) Recognition

RNA fold (35) is a web server for predicting the RNA secondary structure accessible through http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi for anticipation of *S. aurues* tRNA^{phe} secondary structure (Fig.8b). The *T. thermophilus* tRNA^{phe} secondary structure (26) (Fig. 8a) is quite similar to the predicted one for *S. aurues* tRNA^{phe}. In *T. thermophilus,* it was found (28) that G34 interacts with the largest number of amino acid residues as it makes stacking interaction with Tyr β* 731 (*S. aureus* Tyr β* 751) and two base specific interactions. The first one is between O6 G34 and Ser β* 742 (*S. aureus* Ser β* 762) and the second one is between N2 of G34 and Asp β* 729 (*S. aureus* Asp β* 749) in addition to a hydrogen bond between N7 of G34 and Arg β*780 (*S. aureus* Arg β* 800). A35 interacts with Ala β*698 (*S. aureus* Ala β* 718)

through a van der waals bond. A36 interacts with Leu β* 697 (*S. aureus* Ile β* 717) by van der waals and with Asp β*696 (*S. aureus* Asp β* 716) via a hydrogen bond.

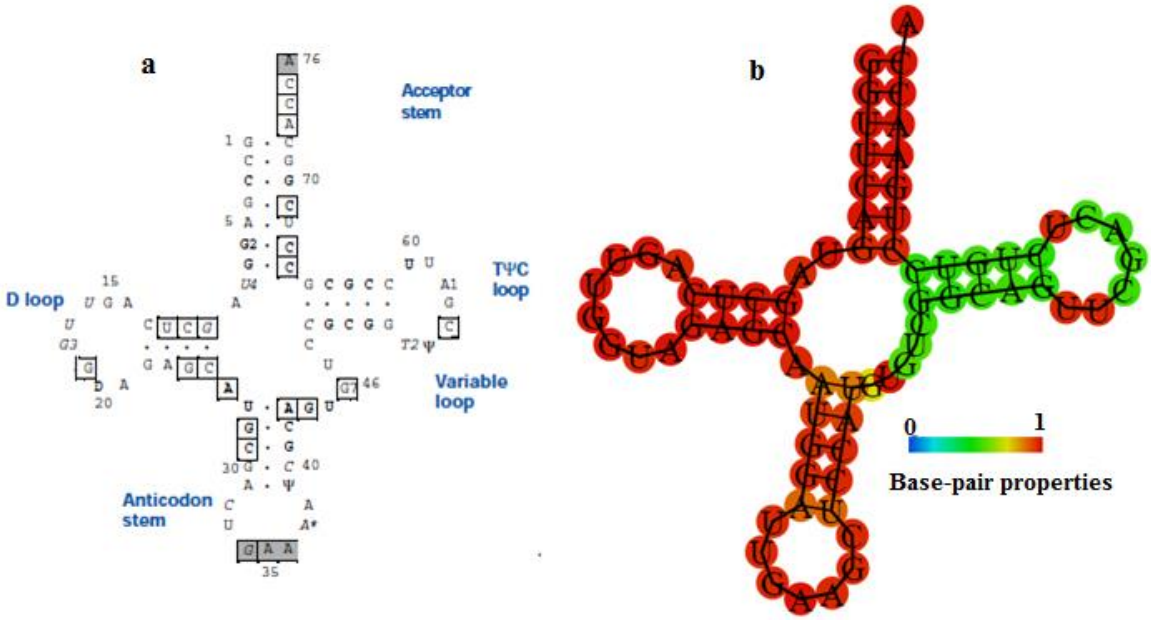


Figure 8 a: Cloverleaf representation of *T. thermophilus* $tRNA^{Phe}$ (26). **b:** predicted secondary structure of the *S.aureus* PheRS $tRNA^{phe}$

Binding with CCA end

According to the ternary complex of *Thermus thermophilus* PheRS.$tRNA^{phe}$.PheOH-AMP (21), the CCA end interacts with the α-subunit active site and with the N-terminal domain of the β-subunit for the same heterodimer. Table 8 shows proposed interactions of the CCA end in $tRNA^{phe}$ with *S. aureus* PheRS based on the clustal alignment of both α and β- subunits for *T. thermophilus* and *S. aureus* and also their genomic alignment of $tRNA^{phe}$.

Table 8 Interactions of the CCA end in both *S. aureus* $tRNA^{phe}$ and *T. thermophilus* $tRNA^{phe}$.

CCA end in *Thermus thermophilus*	Binding residues *T .thermophilus*	CCA end in *Staphylococcus aureus*	Binding residues in *S. aureus*
A76	His α212 (motif 2)	A73	His α208 (motif 2)
	Arg α321 (motif 3)		Arg α318 (motif 3)
	Arg α204 (motif 2)		Arg α200 (motif 2)
	Met α148		Met α157
	Glu α206 (motif 2)		Asp α202 (motif 2)
C75	Glu β31 (β1)	C72	Glu β31 (β1)
	Asp β33 (β1)		Asp β33 (β1)
	His β358 (β3)		Ser β366 (β3)
	Arg β362 (β3)		Lys β370 (β3)
C74	Arg β353 (β3)	C71	Arg β361 (β3)

Final constructed S. aureus PheS model

The final constructed model of *S. aureus* PheS indicates the characteristic domains of aaRSs class II where the active site is built around antiparallel β sheets (Fig.9). The CCA end of $tRNA^{phe}$ interacts with the α-subunit active site and with the N-terminal domain of the β-subunit for the same heterodimer. However, the GAA anticodon $tRNA^{phe}$ is recognised by the β8 domain, which is located at the C-terminal domain of the β*-subunit of the second heterodimer. The model has two metal binding sites, one for magnesium metal, which is present at the interface between the α and β-subunits near to the active site, essential for protein-protein interaction within the PheRS enzyme maintaining its activity in the aminoacylation process. The second one is a zinc finger where Cys267, Cys270, Cys275 and Cys278 make a tetrahedral coordination with zinc (Fig.8).

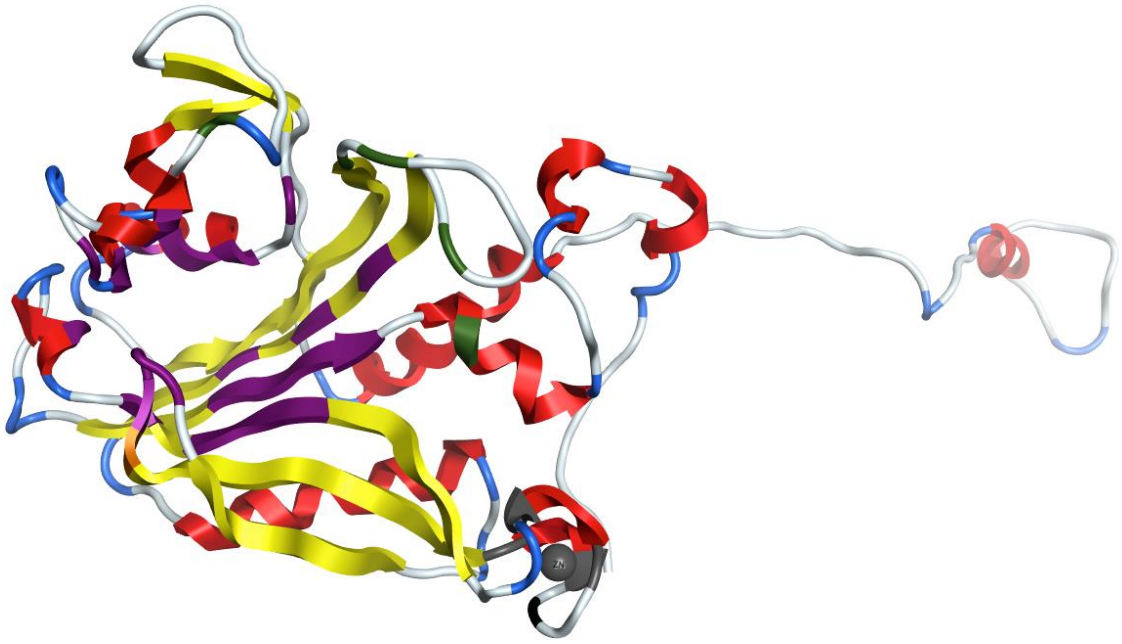


Figure 9 Final *S. aureus* PheS homology model with characteristic domains: α-active site (purple), CCA end of $tRNA^{phe}$ binding site (green), zinc finger (grey), magnesium binding region (orange) and outlier Gly272 (black).

CONCLUSIONS

A homology model for PheRS-α subunit has been developed using a multi-template approach and the constructed model has the characteristic features present in Class II aaRSs enzymes. Docking of phenylalanine, phenylalanyl-adenylate and known inhibitors in the generated model was used to validate the predicted active site through MOE docking procedures. The active site key residues in the model include Ala154, Ser174, Gln214, Glu216, Phe212, Phe256, Glu285 and Gly288. Glu216 is considered as an anchor, directing the ligands into the correct orientation within the hydrophobic pocket. Asp β716, Ile β717, Ala β718, Aspβ749, Tyr β751, Ser β762 and Argβ800 are the predicted residues involved in the specific interactions between the protein and the $tRNA^{phe}$ anticodon (GAA), the corresponding residues in close homologous PheRS enzymes

are conserved at the C terminal of the beta subunit. The CCA end of $tRNA^{phe}$ interacts with the α-subunit active site and with the N-terminal domain of the β-subunit for the same heterodimer. The generated model has two metal binding sites for zinc and magnesium, the later one has an essential role in the binding between the different subunits of the enzyme enabling the enzyme activity. Identification of the whole model with its active binding sites and key binding residues serve as a crucial step in rational drug design for the development of new antibacterial agents.

Acknowledgements

We thank the Egyptian Government for a Channel research scholarship to Samar S. Elbaramawi.

No conflict of interest to declare

REFERENCES:

1. Eriani G, Delarue M, Poch O, Gangloff J, Moras D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature 1990; 347:203–206.
2. Cassels R, Oliva B, Knowles D. Occurrence of the regulatory nucleotides ppGpp and pppGpp following induction of the stringent response in staphylococci. J Bacteriol 1995;177:5161–5165.
3. Cusack S, Hartlein M, Leberman R. Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. Nucleic Acids Res 1991; 19:3489–3498.
4. WHO, Antimicrobial resistance: global report on surveillance, 2014.
5. Hurdle JG, O'Neill AJ, Chopra I. Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents. Antimicrob Agents Chemother 2005; 49: 4821–4833.
6. Keller B, Kast P, Hennecke H. Cloning and sequence analysis of the phenylalanyl-tRNA synthetase genes (pheST) from *Thermus thermophilus*. FEBS Letters 1992; 301: 83-88.
7. Kreutzer R, Kruft V, Bobkova EV, Lavrik OI, Sprinzl M. Structure of the phenylalanyl-tRNA synthetase genes from *Thermus thermophilus* HB8 and their expression in *Escherichia coli*. Nucleic Acids Res 1992; 20: 4173-4178.
8. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 2003; 31: 3784–3788.
9. Savopoulos JW, Hibbs M, Jones EJ, Mensah L, Richardson C, Fosberry A, Downes R, Fox SG, Brown JR, Jenkins O. Identification, cloning, and expression of a functional

phenylalanyl-tRNA Synthetase (pheRS) from *Staphylococcus aureus*. Protein Expres Purif 2001;21: 470–484.

10. Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF. Improving the accuracy of PSI-BLAST protein database searches with composition- based statics and other refinements. *Nucleic acids Res*. **2001**, *29*, 2994-3005.
11. RCSB Protein Data Bank (PDB) http://www.rcsb.org/pdb.
12. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Mi Park Y, Buso N, Lopez R. EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic acids Res 2015; 43: 580-584.
13. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011; 7: 539.
14. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999; 292: 195-202.
15. Molecular Operating Environment (MOE 2014.0901) Chemical Computing Group Inc, Montreal Quebec Canada http://www.chemcomp.com. 2014.0901.
16. Weiner SJ, Kollman PA, Nguyen DT. An all atom forcefield for simulations of proteins and nucleic acids. J Comput Chem 1986;7:230–252.
17. RAMPAGE Server http://ravenbioccam.ac.uk/rampage.php
18. Bowie UJ, Eisenberg D. Method to identify protein sequence that fold into a known three dimensional structure. Science 1991; 253: 164-170.

19. Weiderstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three dimensional structures of proteins. Nucleic acid Res 2007; 35: 407-410.
20. Moor N, Kotik-Kogan O, Tworowski D, Sukhanova M, Safro M. The crystal structure of the ternary complex of phenylalanyl-tRNA synthetase with $tRNA^{Phe}$ and a phenylalanyl-adenylate analogue reveals a conformational switch of the CCA end. Biochemistry 2006; 45:10572-10583.
21. Beyer D, Kroll HP, Endermann R, Schiffer G, Siegel S, Bauser M, Pohlmann J, Brands M, Ziegelbauer K, Haebich D, Eymann C, Bro¨tz-Oesterhelt H. New class of bacterial phenylalanyl-tRNA synthetase inhibitors with high potency and broad-spectrum activity. Antimicrob Agents Chemother 2004; 48: 525–532.
22. Jarvest RL, Erskine SG, Forrest AK, Fosberry AP, Hibbs MJ, Jones JJ, Hanlon P, Sheppard RJ, Worby A. Discovery and optimisation of potent, selective, ethanolamine inhibitors of bacterial phenylalanyl tRNA synthetase. Bioorg Med Chem Lett 2005; 15: 2305–2309.
23. Evdokimov AG, Mekel M, Hutching K, Narasimhan L, et al. Rational protein engineering in action: The first crystal structure of a phenylalanine tRNA synthetase from *Staphylococcus haemolyticus*. J Struct Biol 2008; 162 :152–169.
24. Fishman R, Ankilova V, Moor N, Safro M. Structure at 2.6Å resolution of phenylalanyl-tRNA synthetase complexed with phenylalanyladenylate in the presence of manganese. Acta Crystallogr D Biol Crystallogr 2001; 57: 1534-1544.
25. Finarov I, Moor N, Kessler N, Klipca L, Safro MG. Structure of human cytosolic phenylalanyltRNA synthetase: evidence for kingdom-specific design of the active sites and tRNA binding patterns. Structure 2010; 18: 343-353.

26. Moor N, Lavrik O, Vassylyev DG, Reshetnikova L. Crystal structures of phenylalanyl-tRNA synthetase complexed with phenylalanine and phenylalanyl-adenylate analogue. J Mol Biol 1999; 287: 555-568.
27. Goldgur Y, Mosyak L, Reshetnikova L, Ankilova V, Lavrik O, Khodyreva S, Safro M. The crystal structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus* complexed with cognate $tRNA^{Phe}$. Structure 1997; 5(1):59–68.
28. Mosyak L, Reshetnikova L, Goldgur Y, Delarue M, Safro M. Structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus*. Nat Struct Mol Biol 1995; 2: 537-547.
29. Lechler A, Kreutzer R. The phenylalanyl-tRNA synthetase specifically binds DNA. J Mol Biol 1998; 278: 897-901.
30. Cavarelli J, Rees B, Thierry JC, Moras D. Yeast aspartyl-tRNA synthetase: a structural view of the aminoacylation reaction. Biochimie 1993;75:1117-1123.
31. Nagai AK, Oubridge C, Jessen TH, Li J, Evans PE. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein. Nature 1990; 348:515-520.
32. Lechler A, Kreutzer R. Domains of phenylalanyl-tRNA synthetase from *Thermus thermophilus* required for aminoacylation. FEBS Letters 1997; 420: 139-142.
33. Koch CA, Anderson D, Moran MF, Ellis C, Pawson T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science 1991; 252:668-674.
34. Zuker M, Stiegler P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acid Res 1981; : 133-148.