

Structural basis of the interaction between cyclodipeptide synthases and aminoacylated tRNA substrates

GABRIELLE BOURGEOIS,¹ JÉRÔME SEGUIN,² MORGAN BABIN,² MURIEL GONDRIY,² YVES MECHULAM,¹ and EMMANUELLE SCHMITT¹

¹Laboratoire de Biologie Structurale de la Cellule, BIOC, Ecole polytechnique, CNRS, Institut Polytechnique de Paris, 91128 Palaiseau cedex, France

²Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

ABSTRACT

Cyclodipeptide synthases (CDPSs) catalyze the synthesis of various cyclodipeptides by using two aminoacyl-tRNA (aa-tRNA) substrates in a sequential mechanism. Here, we studied binding of phenylalanyl-tRNA^{Phe} to the CDPS from *Candidatus Glomeribacter gigasporarum* (Cglo-CDPS) by gel filtration and electrophoretic mobility shift assay. We determined the crystal structure of the Cglo-CDPS:Phe-tRNA^{Phe} complex to 5 Å resolution and further studied it in solution using small-angle X-ray scattering (SAXS). The data show that the major groove of the acceptor stem of the aa-tRNA interacts with the enzyme through the basic β 2 and β 7 strands of CDPSs belonging to the XYP subfamily. A bending of the CCA extremity enables the amino acid moiety to be positioned in the P1 pocket while the terminal A76 adenosine occupies the P2 pocket. Such a positioning indicates that the present structure illustrates the binding of the first aa-tRNA. In cells, CDPSs and the elongation factor EF-Tu share aminoacylated tRNAs as substrates. The present study shows that CDPSs and EF-Tu interact with opposite sides of tRNA. This may explain how CDPSs hijack aa-tRNAs from canonical ribosomal protein synthesis.

Keywords: transfer RNA; Rossmann fold; aminoacyl-tRNA synthetases; cyclodipeptides; nonribosomal peptide synthesis

INTRODUCTION

Cyclodipeptide synthases (CDPSs) are enzymes that use sequentially two aminoacyl-tRNAs (aa-tRNAs) to catalyze the formation of two peptide bonds leading to the production of various cyclodipeptides (Fig. 1A; Gondry et al. 2009; Canu et al. 2019). Cyclodipeptides can then be modified by cyclodipeptide-tailoring enzymes in biosynthetic pathways responsible for the production of diketopiperazines with interesting biological activities (Gondry et al. 2009; Aravind et al. 2010; Belin et al. 2012; Borthwick 2012; Giessen and Marahiel 2014; Borgman et al. 2019).

CDPSs are built around a Rossmann fold similar to the catalytic domain of the two class Ic aminoacyl-tRNA synthetases (aaRSs), TyrRS and TrpRS. The catalytic mechanism of CDPSs was extensively investigated (Vetting et al. 2010; Bonnefond et al. 2011; Sauguet et al. 2011; Moutiez et al. 2014a). In particular, the structure of the CDPS from *Streptomyces noursei* (Snou-CDPS, formerly called AlbC) complexed with a reaction intermediate was

determined (Fig. 1B–D; Moutiez et al. 2014a). Overall, the data show that the first aa-tRNA binds the enzyme with its aminoacyl moiety accommodated in a P1 pocket. The aminoacyl moiety is transferred onto a conserved serine residue (S37 in Snou-CDPS) to form an aminoacyl-enzyme intermediate (Sauguet et al. 2011). The second aa-tRNA interacts with the intermediate and its aminoacyl moiety, accommodated in a wide P2 pocket, is transferred to the aminoacyl-enzyme to form a dipeptidyl-enzyme intermediate (Fig. 1D). Finally, the dipeptidyl moiety undergoes an intramolecular cyclization involving a tyrosine (Y202 in Snou-CDPS), properly positioned to favor the nucleophilic attack at the enzyme ester bond, leading to the cyclodipeptide (Moutiez et al. 2014a; Schmitt et al. 2018). In addition to S37 and Y202, residues Y178 and E182 (Snou-CDPS numbering) participate in the catalysis. Y178 and E182 are involved in anchoring the amino and carboxyl groups of the first substrate during the catalytic cycle. Moreover, E182 would act as a catalytic base essential for dipeptidyl-enzyme formation. Notably, the catalytic

Corresponding author: emmanuelle.schmitt@polytechnique.edu

Article is online at <http://www.majournal.org/cgi/doi/10.1261/rna.075184.120>. Freely available online through the RNA Open Access option.

© 2020 Bourgeois et al. This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

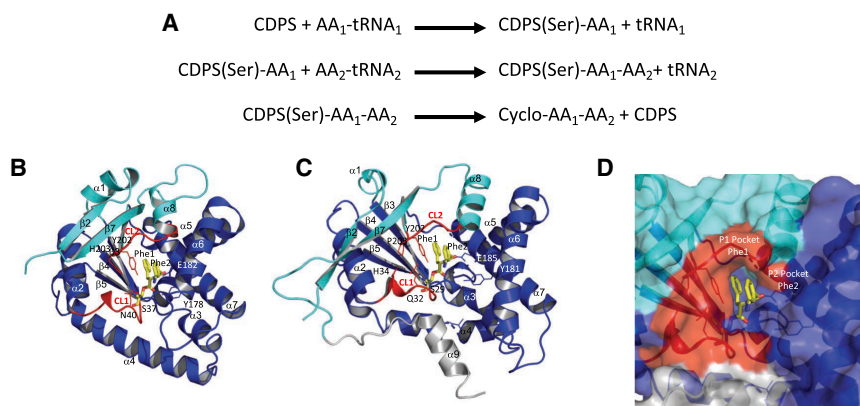


FIGURE 1. Catalytic cycle and binding of the dipeptide to CDPS. (A) Catalytic cycle of CDPS. The aminoacyl-enzyme intermediate is bound to a serine residue. (B) Structure of *Snou*-CDPS (formerly called AlbC) bound to a Phe-Phe dipeptide. The model is deduced from the structure of *Snou*-CDPS bound to a reaction intermediate (PDB ID code 4Q24) (Moutiez et al. 2014a; Schmitt et al. 2018). The first part of the Rossmann fold is colored in cyan and the second part is in dark blue. The two catalytic loops CL1 and CL2 are in red. Residues important for catalysis are shown in sticks. The dipeptide bound to the catalytic serine is in yellow stick. *Snou*-CDPS is a member of the NYH-CDPS subfamily (Jacques et al. 2015; Bourgeois et al. 2018). (C) Structure of *Rgry*-CDPS bound to a Phe-Phe dipeptide. The same color code as in view B is used. *Rgry*-CDPS is a member of the XYP-CDPS subfamily. (D) Closeup of the *Rgry*-CDPS dipeptide binding site from view C. The molecular surface is represented. The view shows the two substrate binding pockets named P1 and P2 (see text).

residues S37, Y178, E182, and Y202 are almost strictly conserved in all active CDPSs (Fig 1B,C; Jacques et al. 2015).

CDPSs are divided into two phylogenetically distinct subfamilies named NYH-CDPSs and XYP-CDPSs according to the occurrence of two sets of conserved residues (Fig. 1B,C; Jacques et al. 2015). Determination of crystal structures of four NYH-CDPSs and three XYP-CDPSs made it possible to give structural bases for the partition of the CDPSs into the two subfamilies (Bourgeois et al. 2018). XYP-CDPSs and NYH-CDPSs mainly differ in the first half of their Rossmann fold. However, the catalytic residues adopt similar positioning regardless of the subfamily (Fig. 1B,C; Bourgeois et al. 2018). Despite these data, interaction between CDPSs and their two aa-tRNA substrates remains poorly documented. Previous work on the NYH *Snou*-CDPS showed that the enzyme discriminates between the two aa-tRNA substrates and possess a specific binding site for each. The binding of the first substrate is strongly dependent on the identity of the aminoacyl moiety of the aa-tRNA whereas both the aminoacyl moiety and the tRNA sequence are essential for the specific recognition of the second substrate (Moutiez et al. 2014b). Nevertheless, no structural data describing a complex between a CDPS and an aa-tRNA was available.

Here, we study the binding of *E. coli* phenylalanyl-tRNA^{Phe} (Phe-tRNA^{Phe}) to the XYP-CDPS from *Candidatus Glomeribacter gigasporarum* (*Cglo*-CDPS) by gel filtration and electrophoretic mobility shift assay. We determine the crystal structure of the *Cglo*-CDPS:Phe-

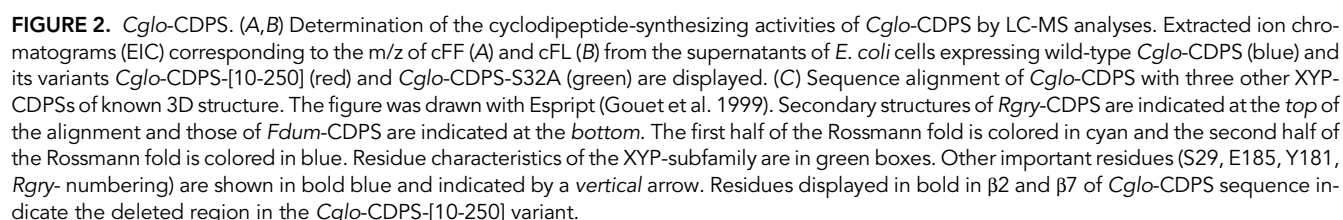
tRNA^{Phe} complex to 5 Å resolution and further validate it by solution studies using small-angle X-ray scattering (SAXS). Despite the modest resolution, structural homologies with other enzymes from the XYP subfamily known at high-resolution allowed us to deduce more informative models for the binding of the aa-tRNA. The data show that the aa-tRNA interacts with the enzyme through the basic β2 and β7 strands of the XYP-CDPS subfamily. A bending of the CCA extremity enables the aminoacyl moiety to be positioned in the P1 pocket while the terminal A76 adenosine occupies the P2 pocket. In cells, CDPSs and the elongation factor EF-Tu (also called EF1A in bacteria, IUBMB 1996) share aminoacylated tRNAs substrates. The present study shows that CDPSs and EF-Tu interact with opposite sides of tRNA. This may explain how CDPSs hijack aa-tRNAs from canonical ribosomal protein synthesis.

RESULTS

Characterization of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex

We used *Cglo*-CDPS to study aa-tRNA binding. The ability of *Cglo*-CDPS to produce cyclo(L-Phe-L-Phe) (cFF) and cyclo(L-Phe-L-Leu) (cFL) in *E. coli* was first measured in Jacques et al. (2015). Here, we confirmed the enzymatic activity of *Cglo*-CDPS in *E. coli*. Moreover, we showed that the replacement of the conserved catalytic serine residue S32 by an alanine (*Cglo*-CDPS-S32A) inactivated the enzyme (Fig. 2A,B). As previously demonstrated for many CDPSs, this mutation impairs the formation of the aminoacyl-enzyme at the first step of catalysis (Vetting et al. 2010; Bonnefond et al. 2011; Sauguet et al. 2011; Moutiez et al. 2014a; Bourgeois et al. 2018). Therefore, this variant allows formation of a CDPS:aa-tRNA complex without further reaction.

E. coli Phe-tRNA^{Phe} (Supplemental Fig. S1A) was aminoacylated using *E. coli* PheRS and then purified (Materials and Methods). In order to test for complex formation, 1.2 molar excess of aminoacylated or nonaminoacylated tRNA^{Phe} were mixed with *Cglo*-CDPS-S32A variant. After 15 min of incubation at 4°C, the complexes were analyzed by size exclusion chromatography (Fig. 3A; Supplemental Fig. S2A). Eluted fractions were analyzed by SDS-PAGE and gels were successively stained with ethidium bromide and Coomassie blue to reveal tRNA and *Cglo*-CDPS,



that a dynamic equilibrium between Cglo-CDPS-S32A and tRNA^{Phe} occurred during the chromatography and therefore that the complex formed with nonaminoacylated tRNA^{Phe} was less stable than that formed with Phe-tRNA^{Phe}

We then used Electrophoretic Mobility Shift Assay (EMSA) to evaluate tRNA binding affinity to the enzyme. Using *Cgl*-CDPS-S32A and Phe-tRNA^{Phe}, we observed the gradual appearance of a shifted band when increasing enzyme concentrations were used (Fig. 4A). As revealed by SYBR gold (Invitrogen) and Coomassie blue staining, this band contained both *Cgl*-CDPS-S32A and Phe-

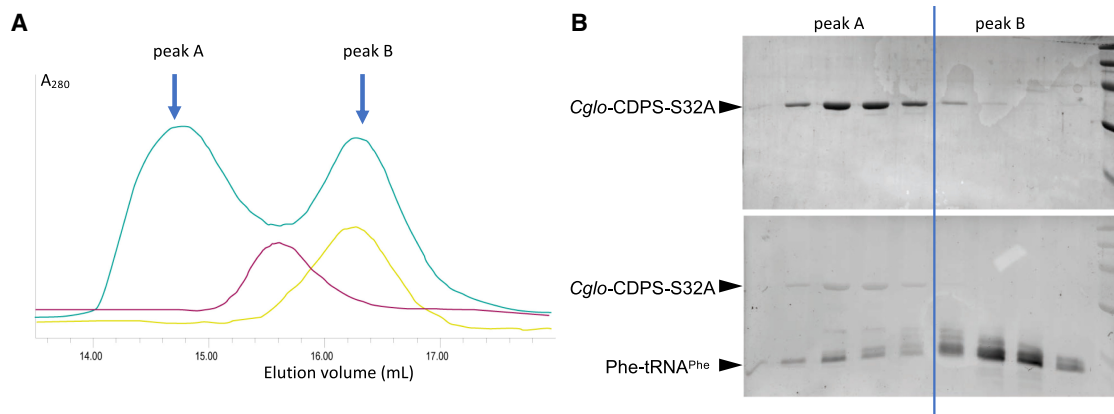


FIGURE 3. Gel filtration analysis of *Cglo*-CDPS:aa-tRNA mixtures. (A) Gel filtration chromatograms for *Cglo*-CDPS-S32A (pink), Phe-tRNA^{Phe} (yellow), *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} mixture (green). (B) SDS-PAGE analysis of fractions from peak A and peak B. Gels were stained using ethidium bromide (lower part) and then Coomassie blue (upper part). The molecular weight marker (GE-Healthcare) is shown to the right of the gel. Visible bands correspond to proteins with molecular weights of 97, 66, 45, 30, 20 kDa from top to bottom. Experimental conditions are described in the Materials and Methods section.

tRNA^{Phe}. An apparent K_D value of $1.6 \pm 0.5 \mu\text{M}$ was deduced from three independent titration experiments. When nonaminoacylated tRNA^{Phe} was used significant dissociation of the CDPS:RNA complex occurs during electrophoresis (Fig. 4B). Therefore an accurate K_D value could not be derived. However, comparison of the EMSA gels (Fig. 4A,B) indicated that apparent tRNA binding affinity for *Cglo*-CDPS-S32A was smaller for the nonaminoacylated form, in agreement with the size-exclusion chromatography experiments. This result pointed out the importance of the phenylalanyl moiety of Phe-tRNA^{Phe} for binding to *Cglo*-CDPS-S32A. Finally, to further investigate if the method allowed us to evidence specific binding of Phe-tRNA^{Phe} to *Cglo*-CDPS-S32A, we used Met-tRNA^{Met}_f^{G1-C72} and nonaminoacylated tRNA^{Met}_f^{G1-C72} (Fig. 4C,D; Supplemental Fig. S1B). This tRNA was used because it is highly overproduced in *E. coli* and easy to purify. It contains a G₁-C₇₂ base pair instead of the A₁-U₇₂ original one to make it an elongator tRNAs (Guillon et al. 1993). As shown in Figure 4C,D, no shifted tRNA band was visible even at high enzyme concentration. Overall the results indicate that *Cglo*-CDPS-S32A specifically binds Phe-tRNA^{Phe} with a K_D value in the micromolar range.

Structure of overproduced *E. coli* tRNA^{Phe}

Cglo-CDPS-S32A:Phe-tRNA^{Phe} complex purified by gel filtration (peak A in Fig. 3A) was used for crystallization using standard sitting drop screening (Hampton research). However, during the search for crystals containing *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex, we identified crystals containing only tRNA^{Phe} (Table 1). In a first step, we used these crystals to determine the structure of tRNA^{Phe} overproduced in *E. coli* at 3.1 Å resolution (Table 1, Materials and Methods). The structure of the overproduced *E. coli*

tRNA^{Phe} contains nucleotides 1 to 75 (Supplemental Fig. S3). Among the 10 modified nucleosides of mature *E. coli* tRNA^{Phe} (Juhling et al. 2009), 3-(3-amino-3-carboxypropyl) uridine (acp³U) was visible at position 47 (Supplemental Fig. S3B). 4-thiouridine (S⁴U) at position 8, dihydrouridine (D) at positions 16 and 20, pseudouridine (ψ) at positions 32, 39, and 55 were also modeled because these modifications are isosteric with the original bases although, at 3.1 Å resolution, no further validation was available in the electron density map. Ribothymidine 54 present in most tRNAs was also modeled. Finally, 2-methylthio-*N*⁶-isopentenyladenosine (ms²i⁶A) at position 37, 7-methylguanosine (m⁷G) at position 46 were not modeled because the modifications were not visible in the electron density map. Comparison of the tRNA^{Phe} structure with that of unmodified tRNA^{Phe} (Byrne et al. 2010) shows that the two molecules are very close to one another (rmsd = 2.56 Å for 1572 atoms compared). In particular, both tRNAs harbor a G10-C25-G44 triplet base pair. Moreover, A26 is unpaired. U45 is flipped out of the tRNA core and forms Watson-Crick interactions with A26 of a symmetry-related molecule (Supplemental Fig. S3C). This packing interaction is notable since the present structure and that of unmodified tRNA^{Phe} were determined from different crystalline forms. Hence, in tRNA^{Phe}, U45 may have a marked propensity to adopt the flipped-out conformation.

Crystal structure of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex

A crystalline form diffracting to 5 Å resolution and containing the *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex was obtained (Table 1). The structure was solved by molecular replacement using two search ensembles; one corresponding to the core domain of *Rgry*-CDPS (Bourgeois

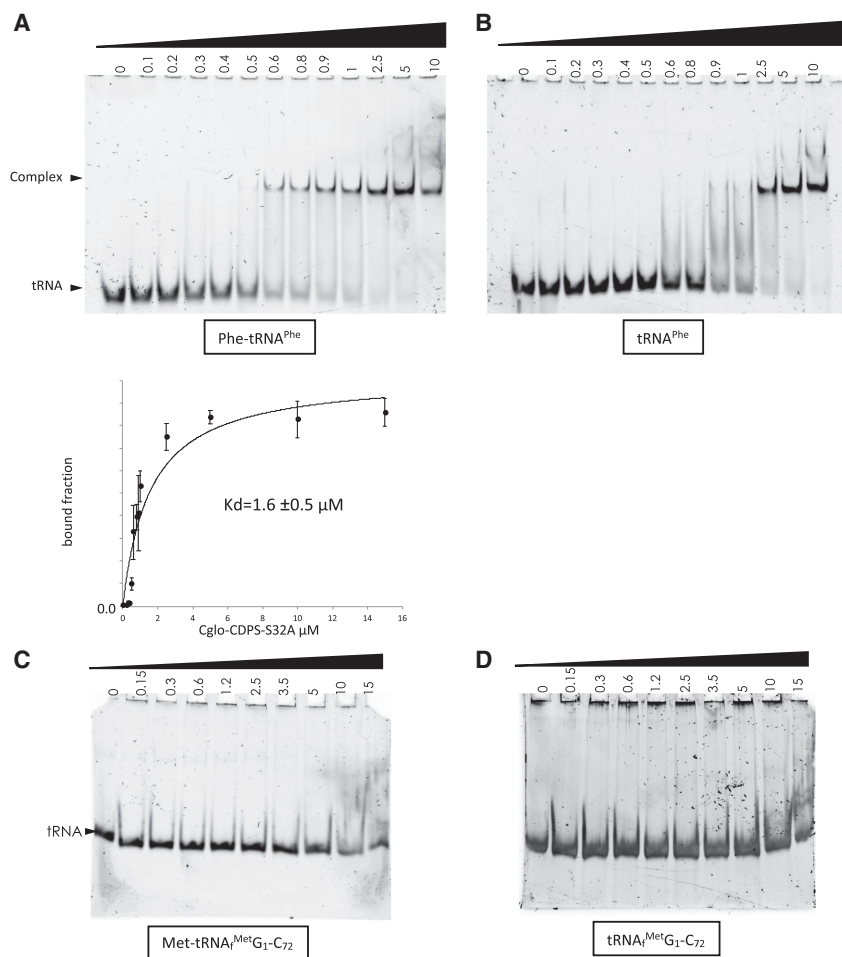


FIGURE 4. Electrophoretic mobility shift assays of *Cglo*-CDPS-S32A:tRNA mixtures. (A) *Cglo*-CDPS-S32A was incubated at increasing concentrations (0.1–15 μ M) with 75 nM Phe-tRNA^{Phe} and electrophoresed on a native acrylamide gel. The panels show typical experiments. All gels were stained using SYBR gold (Invitrogen) and then Coomassie blue. The curve below the gel corresponds to a plot of the intensity of the bound fraction as a function of enzyme concentration. It was fitted according to a simple binding equilibrium. A K_D value of 1.6 ± 0.5 μ M was deduced from three independent experiments. (B) Same experiment as that in A, with 75 nM of tRNA^{Phe}. No K_D value was derived because tRNA dissociation from CDPS is visible during migration. (C) Same experiment as that in A, with 75 nM Met-tRNA^{Met} G1-C72. (D) Same experiment as that in A, with 75 nM tRNA^{Met}.

et al. 2018); residues 12 to 223, 28% identity with *Cglo*-CDPS, Fig. 2D) and the other one to residues 2 to 71 of the structure of *E. coli* tRNA^{Phe} described above. The quality of the molecular replacement density map was sufficient to adjust the models of the tRNA and of the backbone of the protein. Bases 72–75 of the tRNA were then positioned. To avoid overfitting, the protein was refined as a polyaniline model with only a few side chains modeled for residues visible in the electron density. In the final steps of the refinement procedure, weak residual density was visible in the active site at the position expected for the phenylalanyl moiety (Moutiez et al. 2014a). We therefore decided to construct a tentative model of the phenylalanylated A76 base using the structures of the

Snou-CDPS (formerly called AlbC) bound to the analog mimicking the putative dipeptidyl intermediate, ZPCK (N-carbobenzyloxy-L-Phe-chloromethylketone, Moutiez et al. 2014a) and that of a Phe-tRNA^{Phe} molecule coming from a Phe-tRNA^{Phe}:EF-Tu complex (Nissen et al. 1995) as guides. The final model was refined in Phenix (Afonine et al. 2012) to 5 Å resolution (Table 1).

Architecture of the *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex

The 5.0 Å resolution structure of the *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex shows a Phe-tRNA^{Phe} molecule bound to a monomer of *Cglo*-CDPS. Consistent with the previously determined CDPS structures, *Cglo*-CDPS comprises a Rossmann fold with the active site located at the switch point between the two halves of the domain. Moreover, as evidenced by sequence alignments and structure superimposition, the first half of the Rossmann fold is characteristic of enzymes from the XYP subfamily with residues important for catalysis similarly positioned (Figs. 1A,B, 2C; Bourgeois et al. 2018). As compared to other XYP-CDPS sequences, *Cglo*-CDPS has two insertions: one of six residues downstream from β 3 [34–40] and one of 10 residues downstream from α 2 (Fig. 2C). The structure shows that *Cglo*-CDPS approaches Phe-tRNA^{Phe} from the major groove of the acceptor stem via the β 2 and β 7 strands of the first half of the Rossmann fold (Fig. 5A,B). The CCA arm is bent and located in a channel delineated by the two catalytic loops CL1 and CL2 (Fig. 5B). This contributes to the positioning of the aminoacyl group in the catalytic pocket of the enzyme.

Comparison of the structure of free tRNA^{Phe} with that of Phe-tRNA^{Phe} bound to *Cglo*-CDPS (Fig. 5C; rmsd = 3.8 Å for 1611 atoms compared), highlights the bending of the aminoacylated CCA end of Phe-tRNA^{Phe} that accompanies the positioning of the aminoacylated terminal adenosine 76 in the catalytic pocket. As illustrated in Supplemental Figure S4, other parts of the tRNA do not interact with the enzyme but are involved in crystal packing. Indeed, stacking interactions between tRNA molecules involve G34 from the anticodon and the G19:C56 base pair at

Comparison of the structure of free tRNA^{Phe} with that of Phe-tRNA^{Phe} bound to *Cglo*-CDPS (Fig. 5C; rmsd = 3.8 Å for 1611 atoms compared), highlights the bending of the aminoacylated CCA end of Phe-tRNA^{Phe} that accompanies the positioning of the aminoacylated terminal adenosine 76 in the catalytic pocket. As illustrated in Supplemental Figure S4, other parts of the tRNA do not interact with the enzyme but are involved in crystal packing. Indeed, stacking interactions between tRNA molecules involve G34 from the anticodon and the G19:C56 base pair at

TABLE 1. Data collection and refinement statistics

	tRNA ^{Phe}	Cglo-CDPS-S32A:Phe-tRNA ^{Phe}
Data collection		
Molecule in asymmetric unit	1	1
Crystallization conditions	12.4% PEG 8000; 0.2 M calcium acetate; 0.1 M Hepes pH = 7; 0.1 M L-Pro	10% PEG8000; 0.17 M calcium acetate, 0.17 M guanidium hydrochloride, 0.1 M Hepes pH 7
Space group	P6 ₄ 22	P6 ₁ 22
Cell dimensions		
a, b, c (Å)	109.6 109.6 138.5	254.8 254.8 69.4
α, β, γ (°)	90 90 120	90 90 120
Total no. of reflections	117,899 (18,915)	202,634 (28,023)
No. of unique reflections	9432 (1464)	6120 (867)
Resolution (Å)	50–3.1	50–5.0
R _{means} (%)	11.7 (208.9)	17.7 (428.0)
I/σ(I)	14.2 (1.1)	10.7 (1.0)
Completeness (%)	99.7 (98.9)	95.6 (91.6)
Redundancy	12.5 (12.9)	33.2 (32)
CC _{1/2}	99.6 (68.4)	99.7 (58.1)
Data with anisotropic correction (STARANISO server)		
Ellipsoidal resolution (Å)		5.903 (0.894a ^a –0.447b ^a) 5.903 (b ^a) 4.545 (c ^a)
Total no. of reflections (ellipsoidal)		146,988 (10,814)
No. of unique reflections (ellipsoidal)		4373 (337)
R _{means} (%)		16.5 (269.8)
I/σ (I)		15.8 (1.9)
Completeness ellipsoidal (%)		92.7 (93.1)
Completeness spherical (%)		71.5 (25.4)
Redundancy		33.6 (32.1)
CC _{1/2}		100.0 (82.7)
Refinement		
Resolution (Å)	42.9–3.1	48.1–5.0
No. reflections	9420	4315
R _{work} /R _{free}	0.2153/0.2361	0.3015/0.3148
No. atoms	1640	2967
Protein	0	1333
Nucleic acid	1609	1634
Waters	2	0
Heterogen atoms	29	0
B-factors (Å ²) protein	/	112.6
Nucleic acid	137.9	161.9
Water	81.4	/
Heterogen atoms	132.3	/
Root-mean-squared deviation		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.85	0.83

^aValues in parentheses are for highest-resolution shell.

$$R_{\text{sym}}(I) = \frac{\sum_{hkl} \sum_i |I_{hkl} - I_{hkl,i}|}{\sum_{hkl} \sum_i I_{hkl,i}}, \text{ where } i \text{ is the number of reflections } hkl.$$

CC_{1/2} is the correlation coefficient between two random half data sets (Karplus and Diederichs 2012).

$$R_{\text{work}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}.$$

R_{free} is calculated with 5% of the reflections.

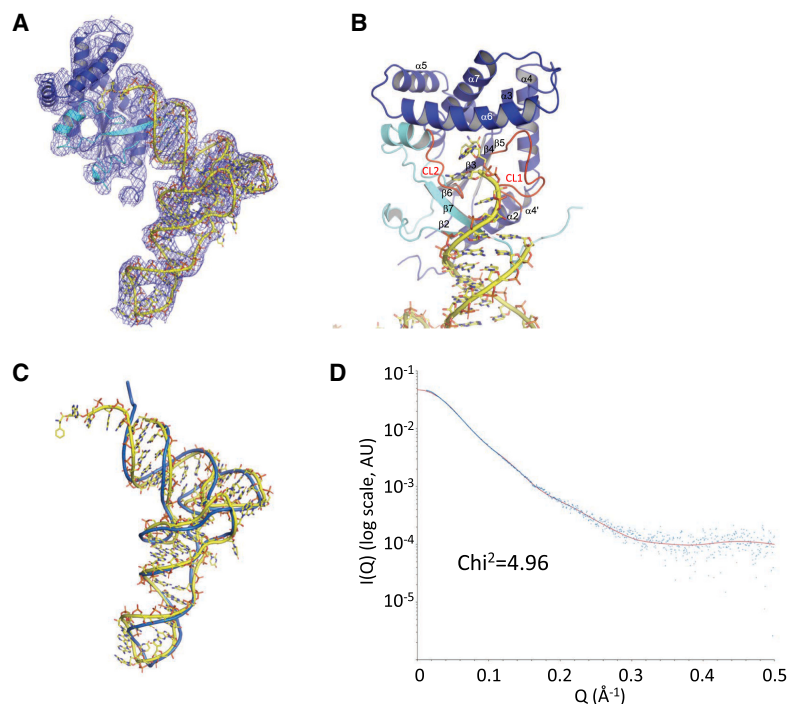


FIGURE 5. Structure of the *Cglo*-CDPS:Phe-tRNA^{Phe} complex. (A) *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex. The first half of the Rossmann fold is colored in cyan and the second half is in dark blue. The $2mF_o - DF_c$ electron density map is contoured at 1 sigma. (B) Closeup of the catalytic center. The two catalytic loops CL1 and CL2 (see Fig. 1) are colored in red. (C) Superimposition of free tRNA^{Phe} onto Phe-tRNA^{Phe} in the *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex. (D) The experimental SAXS curve (blue) is compared with the theoretical diffusion curve deduced from the crystal structure of the *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex (red). AU, arbitrary units. See also Supplemental Figure S6.

the elbow. Moreover, the backbone of the variable loop of tRNA (G44 to U47) contacts the minor groove at the top of the anticodon helix of a symmetry-related molecule. Interestingly, a Watson–Crick interaction between U45 of one molecule and A26 of another molecule stabilizes this interaction, as also observed in the crystals containing only tRNA^{Phe} (Supplemental Fig. S3). Concerning the protein, packing interactions involve the $\alpha 6$ helix on one side and the αA helix on the other side (Supplemental Fig. S4). It should be noted that the acceptor helix of the tRNA is not involved in packing interactions. This makes it unlikely that crystal packing has influenced the architecture of the complex.

Because crystals diffracted to 5 Å resolution, the quality of the electron density did not allow us to build a full atomic model of *Cglo*-CDPS and some parts of the structure of the protein are missing in the model. Therefore, in order to better understand the interaction of tRNA with CDPS, we superimposed the structure of *Rgry*-CDPS (PDB: 5MLP, Bourgeois et al. 2018) onto that of *Cglo*-CDPS (rmsd of 2.03 Å for 194 C α atoms compared). *Rgry*-CDPS is, like *Cglo*-CDPS, a member of the XYP subfamily that produces cFF and cFL (Jacques et al. 2015). Its structure is known at 2.0 Å resolution (Bourgeois et al. 2018). From

the superimposition, we deduced a model of *Rgry*-CDPS bound to Phe-tRNA^{Phe} (Fig. 6). The model was further improved using geometry idealization procedures in phenix.dynamics (Afonine et al. 2012). This model shows that $\beta 2$ and $\beta 7$ strands are embedded into the major groove of the acceptor helix (Fig. 6B,C). The length and basic character of $\beta 2$ and $\beta 7$ strands are known to be features of the XYP-CDPS subfamily (Supplemental Fig. S5A; Bourgeois et al. 2018). These two strands contain basic stretches including a K ϕ X ϕ consensus in $\beta 2$ (where ϕ stands for hydrophobic residue) and a L(R/K)FKK consensus in $\beta 7$ (Bourgeois et al. 2018). Consistent with these findings, the electrostatic potential of the protein highlights the basic character of the $\beta 2$ and $\beta 7$ strands that interact with the acceptor stem of the tRNA (Fig. 6B; Supplemental Fig. S5B). The positioning of the $\beta 2$ and $\beta 7$ strands may help to stabilize the bent conformation of the aminoacylated CCA end of Phe-tRNA^{Phe} (Figs. 5A,B, 6B,C).

In the *Rgry*-CDPS:Phe-tRNA^{Phe} model, the aminoacylated terminal A76 is located within the catalytic

pocket. CDPSs possess two binding pockets, P1 and P2, accommodating the aminoacyl moieties of the first and second aa-tRNA, respectively (Moutiez et al. 2014a; Jacques et al. 2015; Bourgeois et al. 2018). Here, the phenylalanyl moiety occupies the P1 pocket near the residues known to be important for catalysis (S29,Y181,E185,Y202; Fig. 6C) while the adenosine moiety occupies the P2 pocket. The fact that the P2 pocket is wide and shallow likely explains why it is less specific for the amino acid and how it can accommodate the adenosine moiety of the first aa-tRNA substrate. Such a positioning of the aminoacylated part of the tRNA indicates that the present structure illustrates the binding of the first aa-tRNA to the enzyme.

Solution studies using SAXS and site-directed mutagenesis

To further reinforce our data, and in particular to insure that the observed complex between *Cglo*-CDPS and tRNA was not the result of crystallization artifacts, the structure of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} was also studied in solution using SAXS coupled with size-exclusion chromatography (see Materials and Methods and Supplemental Fig. S6). Molecular mass of the complex in solution was

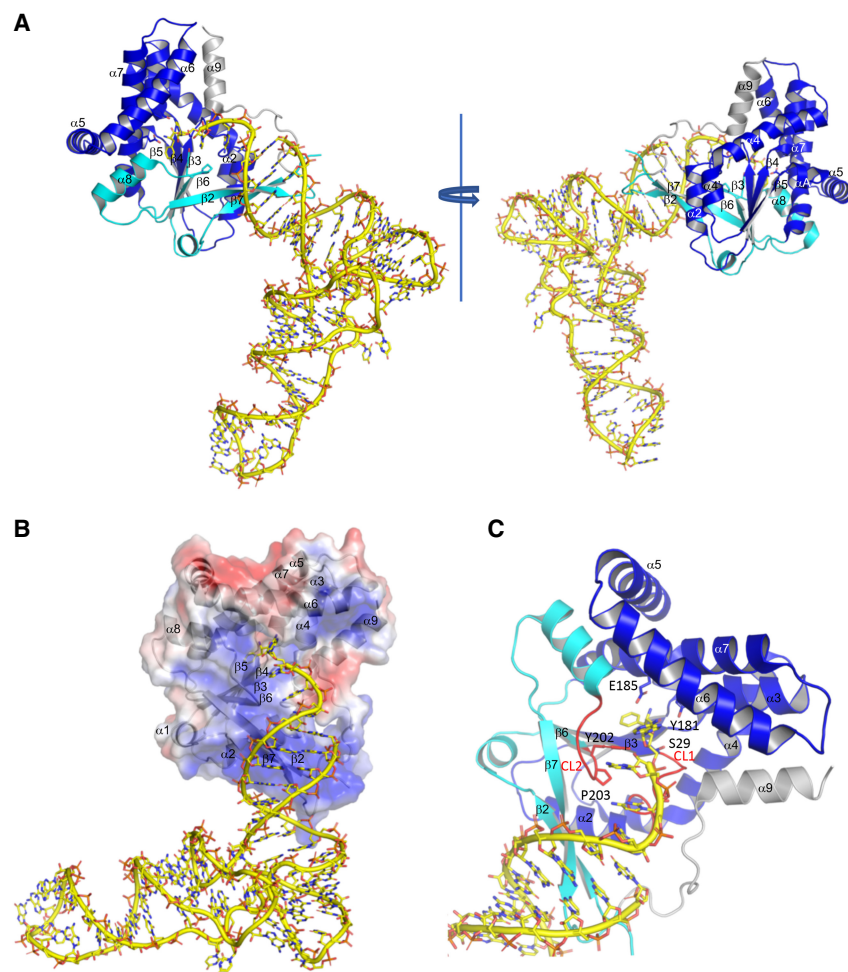


FIGURE 6. Model of the *Rgry*-CDPS:Phe-tRNA^{Phe} complex. (A) The structure of *Rgry*-CDPS refined at 2.0 Å resolution (PDB: 5MLP, Bourgeois et al. 2018) was superimposed onto that of *Cglo*-CDPS to deduce a model of *Rgry*-CDPS:Phe-tRNA^{Phe}. The color code is the same as in Figure 5. The carboxy-terminal helix of *Rgry*-CDPS is in gray. The two structures are superimposed with an rmsd of 2.03 Å for 194 Ca atoms compared. (B) Electrostatic potential map of *Rgry*-CDPS (blue, positive; red, negative; white, neutral) with the tRNA shown in yellow sticks and cartoon. The views emphasize the binding of the acceptor arm of the tRNA by the positively charged protein area containing $\beta 2$ – $\beta 7$ (see also Supplemental Fig. S4). The electrostatic potential map was calculated using the APBS plugin in Pymol (Schrodinger 2017). (C) Closeup of the catalytic center in the *Rgry*-CDPS:Phe-tRNA^{Phe} model. Figures 1, 5–7 were drawn with Pymol.

estimated from the SAXS data to 53 ± 3 kDa using the consensus Bayesian approach (Hajizadeh et al. 2018) and to 58 kDa using the MoW method (Piiadov et al. 2019). This agrees with a 1:1 CDPS:tRNA complex (calculated molecular mass of 58 kDa), as observed by gel filtration (data not shown) and in the crystal structure. The experimental SAXS curve measured with the size-exclusion chromatography-purified complex was then compared with the curve calculated from the crystal structure coordinates (Fig. 5D). The overall agreement was good ($X^2 = 4.96$) arguing in favor of the biological significance of our crystalline model. Small deviations between the two curves were however observed, in particular at low Q values where experimental

diffusion is larger than expected. When not bound to tRNA, *Cglo*-CDPS tends to form aggregates that diffuse X rays at low Q (SAXS data not shown). Thus, one hypothesis to account for the small deviation of the two curves at low Q may be a very small amount of dissociation of the complex followed by interactions between the liberated protein molecules, in the path between the chromatography column and the SAXS cell.

To further validate the involvement of the $\beta 2$ and $\beta 7$ strands in the binding of Phe-tRNA^{Phe}, we designed variants of *Cglo*-CDPS and *Cglo*-CDPS-S32A in which residues 1 to 9 and residues 251–293 were removed (see bold characters in $\beta 2$ and $\beta 7$ in *Cglo* sequence Fig. 2C). The variants were named *Cglo*-CDPS-[10-250] and *Cglo*-CDPS-S32A-[10-250]. First, cyclodipeptide-synthesizing activity was measured from overproducing cells as described in the Materials and Methods section. The presence of the truncated protein in the soluble fraction of crude extracts was confirmed by SDS-PAGE analysis (Supplemental Fig. S7A,B) but no cyclodipeptide-synthesizing activity was detected for the *Cglo*-CDPS-[10-250] variant (Fig. 2A,B). The truncated protein was purified and its correct folding was checked by circular dichroism (Supplemental Fig. S7C). In a second step, the ability of the *Cglo*-CDPS-S32A-[10-250] variant to bind Phe-tRNA^{Phe} was analyzed by gel filtration (Supplemental Fig. S7D). When aminoacylated tRNA^{Phe} and *Cglo*-CDPS-S32A-[10-250] variant were

mixed using a 1.2:1 molar ratio, only one large peak with a maximum corresponding to the elution peak Phe-tRNA^{Phe} was observed, arguing in favor of the absence of complex formation (Supplemental Fig. S7D). This confirmed the importance of the $\beta 2$ and $\beta 7$ strands for tRNA binding, as indicated by the crystallographic structure.

DISCUSSION

This work describes the first crystallographic structure of a *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex at 5 Å resolution. The binding mode allows positioning of the esterified phenylalanine within the catalytic pocket close to the catalytic

serine residue. Therefore, the structure illustrates the binding of the first aa-tRNA. Higher affinity of Phe-tRNA^{Phe} for the first site is at least partly explained by complementarity of the phenylalanyl moiety with P1. After reaction, Phe acylated to the catalytic serine will contribute to shape the otherwise wide P2 pocket thereby favoring binding of the second aa-tRNA.

The structure shows that the β 2 and β 7 strands of the first part of the Rossmann fold of *Cglo*-CDPS are involved in tRNA binding. Interestingly, this part of the CDPS structures has been previously shown to provide a structural basis for the partition of the CDPSs into the two XYP and NYH-CDPS subfamilies. The crystallographic structures of three other XYP-CDPSs are known. Among them, *Rgry*-CDPS catalyzes formation of cFF and cFL, like *Cglo*-CDPS. The experimental structure allowed us to readily derive a model for the binding of the tRNA to *Rgry*-CDPS, whose structure is known at high-resolution. This model shows how the β 2 and β 7 strands are embedded into the major groove of the acceptor stem of Phe-tRNA^{Phe}. A docking model performed with all known XYP-CDPS structures highlights the role of the basic β -strands in tRNA selection for this subfamily (Fig. 6B,C; Supplemental Fig. S5B–D). This raises the question of tRNA recognition in the NYH-CDPS subfamily. Given the structural proximity of the two CDPS subfamilies, docking models of tRNA on the NYH-CDPS can also be deduced from the structure of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex (Supplemental Fig. S5E–G). The figure shows that the β 2 and β 7 strands, shorter in NYH-CDPS, do not provide a large and basic interface favorable to tRNA binding. Moreover, a previous study has proposed that the α 4 helix of *Snou*-CDPS could be involved in tRNA binding (Moutiez et al. 2014b). Therefore, it cannot be excluded that the binding of the acceptor stem of the first aa-tRNA is different in the two CDPS subfamilies even if the aminoacylated A76 base would be accommodated in the same cavity. Further experiments and structural data are required to fully address this question.

The structural homology of the CDPSs with class 1 aaRSs has been noted very early, the closest representatives being TyrRS and TrpRS (Vetting et al. 2010; Bonnefond et al. 2011; Sauguet et al. 2011). TyrRS and TrpRS belong to class1c aaRS (Giegé and Springer 2012). In contrast to other class I aaRS members, these two enzymes recognize their cognate tRNAs via the major groove of their acceptor stem. However, the 3' end CCA of the tRNA is bound at the catalytic site with a bent conformation, which is a typical binding mode of class I aaRSs (Rould et al. 1989; Yaremchuk et al. 2002; Kobayashi et al. 2003; Shen et al. 2006; Yang et al. 2006; Carter and Wills 2018). In order to compare the tRNA binding mode of TyrRS and TrpRS with that of XYP-CDPS, we superimposed the Rossmann folds of the enzymes (Fig. 7A; Supplemental Fig. S8). The tRNA binding mode is totally different for CDPS and

for TyrRS or TrpRS. The tyrosine and tryptophane amino acid substrates are located in the same binding pocket as the phenylalanyl moiety of Phe-tRNA in CDPS (P1 pocket). The ATP ligand of the two aaRS is located opposite to the A76 base of Phe-tRNA^{Phe} in the XYP-CDPS:Phe-tRNA^{Phe} complex. Interestingly, structural alignment of the two aaRS with the XYP-CDPS shows that the two enzymes are more divergent in the first part of their Rossmann folds (Fig. 7A; Supplemental Fig. S8). One may imagine that structural variability of this part of the Rossmann fold is linked to tRNA binding specificity. In human TrpRS-tRNA^{Trp}, the 3' end of tRNA^{Trp} bound at the catalytic site with a bent conformation is well defined in the structure (Shen et al. 2006). The tRNA^{Trp} acceptor arm is primarily recognized by an α -helix of the amino-terminal domain and has no base-specific interaction with the catalytic domain. It has been proposed that interaction of the A76 phosphate with the amino terminus of an α -helix is a constant feature facilitating the bent conformation of tRNAs bound to class I aaRS (Hol et al. 1978; Carter and Wills 2018). In contrast to class I aaRS (Carter and Wills 2018), there is no amino terminus of an α -helix stabilizing the conformation of A76 in CDPS. Taking into account the low resolution of the structure, it is not possible to give an atomic description of the CCA arm interactions with the protein. Nevertheless, interactions of β 2 and β 7 with the major groove of the acceptor stem likely favor the bent conformation of the tRNA (Figs. 5B, 6C). In this view, positively charged residues in this region (Fig. 2C) would play an important role. Moreover, the structure shows a channel formed by the two catalytic loops CL1 and CL2 located from either side of the CCA arm that contributes to its orientation toward the catalytic center. Notably, in *Cglo*-CDPS, CL1 loop is extended by seven residues. This extension could bring more surface interaction and thereby favor complex stability. In the case of *Rgry*-CDPS, the carboxy-terminal α 9 helix is packed onto CL1. This interaction could also contribute to stabilize the bent conformation of the CCA arm.

In cells, CDPSs and other peptide bond-forming enzymes utilizing aa-tRNAs as substrates, such as L/F transferases or Fem aa-transferases (Fonvielle et al. 2013), are in competition with EF-Tu for the availability of aa-tRNA substrates. Consistent with this idea, the initial rate of L/F-transferase activity was shown to be reduced in the presence of an excess amount of EF-Tu (Suto et al., 2006). In order to better understand the molecular basis of such a competition, we compared the tRNA binding mode of EF-Tu with that of CDPSs. As shown in Figure 7B, aa-tRNAs are recognized through opposite sides by EF-Tu and CDPSs. Different types of bending of the acceptor ends enable the esterified amino acid to reach its binding pocket on the two proteins. We speculate that this may help CDPSs to hijack an aa-tRNA already bound to EF-Tu (Fig. 7C). Transient flipping of the aminoacylated CCA end toward

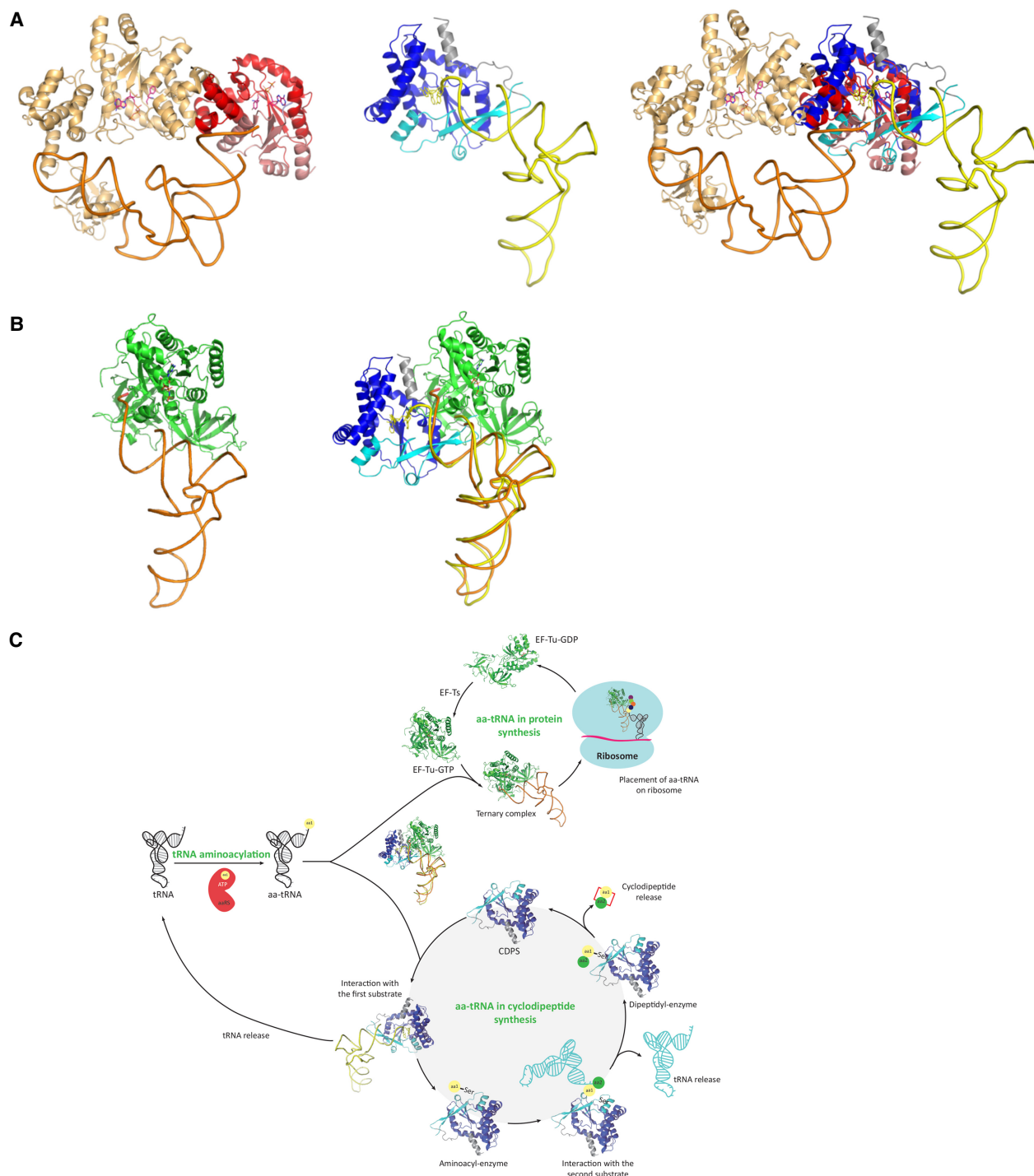


FIGURE 7. Comparison of tRNA binding modes. (A) Structural alignment of *Rgry*-CDPS:Phe-tRNA^{Phe} with *T. thermophilus* TyrRS:tRNA^{Tyr}. *Left* view, the *T. thermophilus* TyrRS:tRNA^{Tyr} complex is represented in cartoon (PDB ID Code 1H3E, Yaremchuk et al. 2002). One monomer is in light orange and the amino-terminal domain of the second monomer is in colored in light red (first part of the Rossmann fold) and in dark red (second part of the Rossmann fold). ATP and tyrosine are shown in sticks. *Middle* view, *Rgry*-CDPS:Phe-tRNA^{Phe} model. Phe-A76 and C75 are shown in yellow sticks. The color code is the same as in Figure 5. *Right* view, superimposition of the two structures. The view shows that the tRNA binding mode is different for *Rgry*-CDPS and for TyrRS even if the two enzymes have similar catalytic domain (the Rossmann fold of the two enzymes superimpose with an rmsd = 3.5 Å for 159 matched Ca positions). (B) Comparison of Phe-tRNA^{Phe} binding to EF-Tu and to *Rgry*-CDPS. *Left* view, *T. thermophilus* EF-Tu:GDPNP:Phe-tRNA^{Phe} complex (PDB ID Code 1TTT, Nissen et al. 1995). *Middle* view, the Phe-tRNA^{Phe} molecules of the EF-Tu:GDPNP:Phe-tRNA^{Phe} and *Rgry*-CDPS:Phe-tRNA^{Phe} complexes were superimposed. The view shows how EF-Tu and *Rgry*-CDPS interact with different parts of the Phe-tRNA^{Phe} molecule. (C) Hijacking of elongator tRNAs by CDPSs. The figure shows that aa-tRNAs are diverted from their canonical role in ribosomal protein synthesis to catalyze the formation of two peptide bonds leading to the production of cyclodipeptides.

the CDPS active site may lead to irreversible transfer of the amino acid to the catalytic serine. Alternatively, transient binding of CDPS to an aa-tRNA already bound to EF-Tu may accelerate its dissociation immediately followed by CDPS binding and catalysis. Interestingly, such an idea has already been proposed for two other enzymes interacting with aminoacylated-tRNA, the methionyl-tRNA^{Met} transformylase (Schmitt et al. 1998) and the L/F-tRNA protein transferase (Suto et al. 2006). It is tempting to imagine that these opposite recognition modes generally contribute to the regulation of cellular processes.

MATERIALS AND METHODS

Bacterial strains and plasmids

E. coli XL1-blue and BL21 Rosetta pLacI-Rare (Invitrogen) have been used for cloning and expression experiments, respectively. The plasmid encoding *Cglo*-CDPS (pIJ196-CDPS23) was described (Jacques et al. 2015). It derives from pQE60 (Qiagen, ColE1 origin and ampicillin resistance) and carries the CDPS gene under the control of an IPTG-inducible PT5 promoter (Jacques et al. 2015). pIJ196-CDPS23 allows expression of the full-length CDPS with the addition of an A residue after the amino-terminal methionine and of the RS sequence followed by a His₆ tag after the last residue.

Site-directed mutagenesis

Thermo Scientific Phusion DNA polymerase was used to perform site-directed mutagenesis of *Cglo*-CDPS gene according to the manufacturer's protocol. Mutation of the catalytic serine into alanine was performed with the following oligonucleotides: 5'-GG CATTGCCATGCAGAGTCCGCATCAGATTAGC-3' and 5'-GGA CTCTGCATGGCAATGCCAATCACGATTTTAATC-3'. Deletions of amino- and carboxy-terminal parts of *Cglo*-CDPS residues were performed using the following oligonucleotides: Δ [1–9] 5'-GAGGAGAAATTAACCATGCTGGAACAGTTCATTCTGG-3' and 5'-CCAGAATGAAGTGTCCAGCATGGTTAATTTCTCCTC-3', Δ [251–293] 5'-CCGCTGCCGTTTATTGAACATCACCATCACCATC ACTAAGC-3' and 5'-GCTTAGTGATGGTGATGGTGATGTTCAA TAAACGGCAGCGG-3'.

Production and purification of *Cglo*-CDPS variants

BL21 Rosetta pLacI-Rare cells were transformed with pIJ196-CDPS23 and grown in 2xTY medium containing 50 μ g/mL ampicillin and 34 μ g/mL chloramphenicol at 37°C until OD_{600nm} = 1. Protein expression was then induced by addition of isopropyl- β -D-thiogalactopyranoside at a final concentration of 0.5 mM at 20°C during 18 h. Cells were resuspended in lysis buffer (10 mM HEPES pH = 7.5, 500 mM NaCl, 3 mM β -mercaptoethanol) and disintegrated by sonication. *Cglo*-CDPS was purified via Talon Metal affinity chromatography (Clontech) using standard protocols followed by anion exchange chromatography (Q sepharose Fast Flow, GE Healthcare). Enzyme concentration was calculated

using a computed extinction coefficient of 33140 M⁻¹cm⁻¹. About 1 mg of *Cglo*-CDPS was obtained from 1 L of culture.

Production of CDPSs and characterization of cyclodipeptide-synthesizing activities

Recombinant CDPSs, wild-type *Cglo*-CDPS and variants *Cglo*-CDPS-S32A and *Cglo*-CDPS-[10-250] were expressed in *E. coli* BL21AI [pREP4] cells, respectively transformed by the plasmid pIJ196-CDPS23 (Jacques et al. 2015), and the corresponding mutated plasmids (see above). The strains were grown in 10 mL M9 minimum medium supplemented with trace elements and vitamins (Gondry et al. 2009), 0.5% glucose, 200 μ g/mL ampicillin and 25 μ g/mL kanamycin. After overnight incubation at 37°C, this starter culture was used to inoculate 200 mL of the same medium except that glucose was replaced by 0.5% glycerol. Bacteria were grown at 37°C until the OD₆₀₀ reached 0.6. The expression of CDPSs was induced by the addition of IPTG (2 mM final concentration) and cultivation was continued for 18 h at 20°C. The cultures were then centrifugated at 4000g for 45 min. Supernatants were analyzed for the presence of cyclodipeptides and CDPS expression was analyzed in cell pellets, as follows.

Supernatants were acidified with trifluoroacetic acid at 2% v/v final, and stored at –20°C. They were then analyzed by LC-MS/MS analyses on an Elute SP HPLC chain (Bruker Daltonik GmbH) coupled via a split system to an amazon SL ion trap mass spectrometer (Bruker Daltonik GmbH) set in positive mode. Samples were loaded onto a C18-PFP column (4.6 mm \times 150 mm, 3 μ m, 100 Å, ACE) developed over 20 min with the linear gradient 0%–50% (v/v) (solvent A: 0.1% [v/v] formic acid in H₂O, solvent B: 0.1% [v/v] formic acid in acetonitrile/H₂O [90/10], flow rate, 0.6 mL/min). Cyclodipeptides were detected and identified as previously described (Fig. 2A,B; Canu et al. 2019).

Cell pellets were frozen at –80°C and treated as described except that they were suspended in only 4 mL of ice-cold lysis buffer (100 mM Tris HCl, pH 8.0, 150 mM NaCl, 5% glycerol [v/v], 10 mM beta-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride) (Gondry et al. 2009). An amount of 550 μ L of the soluble fractions were subjected to spin purification using Micro Bio-Spin chromatography columns filled with 100 μ L of Protino Ni-NTA Agarose resin. The columns were washed with 100 μ L of buffer A containing 80 mM imidazole and the tagged proteins were eluted with 100 μ L buffer A containing 350 mM imidazole. Eluted fractions were analyzed by 12% SDS-PAGE with Coomassie blue staining (Supplemental Fig. S7A).

CD spectra of *Cglo*-CDPS-S32A and *Cglo*-CDPS-[10-250] were recorded using a Jobin-Yvon Mark VI circular dichrograph at a scan speed of 0.2 nm/sec. The protein concentrations were 5 μ M. Quartz cuvettes with 0.1 cm path length per compartment were used. The results are presented as normalized $\Delta\epsilon$ values on the basis of the amino acid mean residue mass of 110 Da (Supplemental Fig. S7C).

Production and aminoacylation of tRNA^{Phe}

A synthetic gene for *E. coli* tRNA^{Phe} was assembled from six overlapping oligonucleotides and cloned into pBST-NAV2 (Meinzel et al. 1992). *E. coli* tRNA^{Phe} was then overexpressed in JM101Tr cells transformed with pBST-NAV2-tRNA^{Phe} in a 1 L 24 h culture

at 37°C. tRNA^{Phe} was purified as described previously for tRNA^{Met} (Mechulam et al. 2007). Briefly, a whole tRNA extract was prepared according to the Zubay procedure (Zubay 1962). The overexpressed tRNA^{Phe} was then purified by an anion exchange step on a Q-Hiload column (16 × 100 mm) equilibrated in 20 mM Tris-HCl pH 7.6, 200 mM NaCl, 8 mM MgCl₂, 0.1 mM EDTA. tRNAs were eluted using a 350 mM to 550 mM NaCl gradient in the same buffer. About 10 mg of tRNA^{Phe} was obtained from 1 L of *E. coli* culture. ¹⁴C Phenylalanine acceptance of the tRNA^{Phe} preparations ranged between 1450 and 1600 pmol/A₂₆₀ unit. tRNA (10 μM) was aminoacylated during 20 min at 25°C using purified *E. coli* PheRS (1 μM) in a buffer containing 10 mM HEPES pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 8 mM MgCl₂, 2 mM ATP and 400 μM Phenylalanine. PheRS was overproduced in *E. coli* IBPC1671 from plasmid pBH16D (Mechulam et al. 1987) and purified using standard chromatographic procedures. After aminoacylation, tRNA was ethanol precipitated in the presence of 166 mM sodium acetate pH 5.0. The aminoacylated tRNA was then purified onto a MonoQ column (GE Healthcare) using a 0 mM to 400 mM NaCl gradient followed by a 800 mM NaCl eluting step. The recovered fractions were ethanol precipitated and the aminoacylated tRNA was finally resuspended in buffer A containing 10 mM MOPS pH 6.7, 200 mM NaCl, 8 mM MgCl₂, and 5 mM 2-mercaptoethanol.

Gel filtration analysis of CDPS:tRNA mixtures

1.2 molar excess of aminoacylated or nonaminoacylated tRNA^{Phe} were mixed with *Cglo*-CDPS-S32A variant in buffer A. The mixtures were kept 15 min at 4°C before analysis by size exclusion chromatography onto a Superdex 200 Increase 10/300 column (0.75 mL/min) equilibrated in buffer A. After gel filtration, fractions were analyzed by SDS-PAGE and gels were successively stained with BET and Coomassie blue to reveal tRNAs and proteins, respectively (Fig. 3; Supplemental Fig. S2).

Electrophoresis mobility shift assay

Electrophoretic analyses were performed on native 12% acrylamide gels in buffer A containing 90 mM Tris pH 8.3, 90 mM sodium borate, 2 mM EDTA and 0.5 mM MgCl₂. tRNA (75 nM) was incubated with increasing concentrations of *Cglo*-CDPS (from 0.1 to 15 μM) for 20 min at 4°C in buffer A. Before loading, 10 μL of the mixture was supplemented with 2 μL loading buffer (0.0025% formamide blue, 0.0025% xylene cyanol, 40% glycerol). The gels were run at 100 V for 2 h at 4°C and then stained successively with SYBR gold (Thermofisher) to reveal the bands containing tRNA and Coomassie blue to reveal the bands containing protein.

Crystallization and structure determination of *E. coli* tRNA^{Phe} and of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex

Freshly gel filtration purified CDPS-S32A:Phe-tRNA^{Phe} complex, as described in Figure 3, was used in vapor diffusion crystallization trials at 4°C using hanging drops obtained by mixing 2 μL of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex (150 μM) with 1.6 μL of precip-

itating solution (Table 1). Two types of crystal were obtained. Before data collection, both crystal types were soaked in a cryoprotecting solution containing the precipitating solution plus 30% (v/v) glycerol. Diffraction data were collected at 100 K at the Proxima 2 beamline (SOLEIL Synchrotron). Diffraction images were analyzed with XDS (Kabsch 1988) and further processed with programs of the CCP4 package (Collaborative Computational Project Number 4, 1994). Both crystal types belonged to a hexagonal space group (Table 1). Crystals containing only tRNA^{Phe} diffracted to 3.1 Å resolution and those containing the CDPS-S32A:Phe-tRNA^{Phe} complex diffracted to 5.0 Å resolution. In a first step, the structure of *E. coli* tRNA^{Phe} was solved by molecular replacement with PHASER (McCoy et al. 2007) using the [1–71] region of *E. coli* tRNA^{Phe} from 4YCO PDB entry (Byrne et al. 2015) as a search model (Z-score = 18.1) and refined with PHENIX (Adams et al. 2010). The final structure of *E. coli* tRNA^{Phe} contains nucleotides 1 to 75.

A highly redundant data set was collected from one crystal containing *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex to 5 Å resolution using the helical collection mode at Proxima 2 (Table 1). After data processing with XDS (Kabsch 1988), the data set was corrected for anisotropy using the STARANISO server (<http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>, Table 1) (Vonnrhein et al. 2018). The structure of *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex was solved by molecular replacement with PHASER (Storoni et al. 2004) using two ensembles. The first one corresponded to the core domain of *Rgry*-CDPS (residues 12–223, PDB:5MLP [Bourgeois et al. 2018]) and the second one to the presently determined 2–71 segment of the structure of *E. coli* tRNA^{Phe}. For each ensemble, a solution with a high Z-score value was found (Z = 12.3 and Z = 6.8). The molecular replacement 2mF_o – DF_c electron density map was of good quality. In particular, additional electron density was visible for nucleotides 72–75 of the tRNA that had been omitted in the molecular replacement model. The secondary structure elements of *Cglo*-CDPS were placed in the electron density. Considering the low diffraction limit, the protein model was built as a polyaniline model with only a few side chains to avoid data overfitting. Coordinates were refined through several cycles of manual adjustment with Coot (Emsley et al. 2010) and positional refinement with PHENIX (Adams et al. 2010) with restraints on secondary structure elements of the protein and of the tRNA. In the last steps of refinement, we modeled the aminoacylated A76 adenosine as described in the Results section. The final model refined to 5 Å resolution contained all nucleotides of Phe-tRNA^{Phe} and residues 9–69 and 75–268 of *Cglo*-CDPS (Table 1).

SAXS data collection

SAXS experiments were conducted at the SWING beamline of the SOLEIL Synchrotron (λ = 1.033 Å). Data were collected in the Q-range 0.008–0.6 Å^{–1} (Q = 4πsinθ/λ, 2θ is the scattering angle). All solutions were circulated in a thermostated (15°C) quartz capillary with a diameter of 1.5 mm and a wall thickness of 10 μm, positioned within a vacuum chamber. For data collection, purified *Cglo*-CDPS-S32A:Phe-tRNA^{Phe} complex (c.a. 20 nanomoles in 100–200 μL) was injected onto a size-exclusion Superdex 200 increase 3.2/300 column (GE Healthcare) using an Agilent High Performance Liquid Chromatography system and eluted directly in buffer A into the SAXS capillary cell at a flow rate of 0.075

mL/min. SAXS data were collected online, with a frame duration of 1.0 sec. A large number of frames were collected during the first minutes of the elution and averaged to account for buffer scattering, which was subsequently subtracted from the signal during elution of the complex. Selected curves corresponding to the main elution peak were averaged on the basis of identical shapes (David and Perez 2009). Data reduction to absolute units, frame averaging and subtraction, were performed using FOXTROT (<http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING>). All subsequent data processing, analysis and modeling steps were carried out with PRIMUS and other programs of the ATSAS suite (Konarev et al. 2006). Scattered intensity curves were calculated from the atomic coordinates of the crystallographic structures using CRY SOL (Svergun et al. 1995) with 50 harmonics. This program was also used to fit the calculated curve to the experimental one, by adjusting the excluded volume, the averaged atomic radius and the contrast of the hydration layer surrounding the particle in solution.

DATA DEPOSITION

Coordinates and structure factors have been deposited in the PDB with accession numbers 6Y3G for *E. coli* overproduced tRNA^{Phe} and 6Y4B for Cglo-CDPS-S32A:Phe-tRNA^{Phe} complex. SAXS data have been deposited at the SASDB (accession code SASDJT4).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by grant ANR-14-CE09-0021 (Agence Nationale de la Recherche). G.B. and M.B. were supported by fellowships from grant ANR-14-CE09-0021. This work was also supported by grants from the Centre National de la Recherche Scientifique and by Ecole polytechnique to Unité Mixte de Recherche no. 7654. We thank the staff of the PX1, PX2, and SWING beamlines at the SOLEIL synchrotron (Saclay, France) for expert assistance during data collection. We thank Magali Blaud (LRCB, Paris) for CD measurements.

Received February 26, 2020; accepted July 6, 2020.

REFERENCES

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**: 213–221. doi:10.1107/S0907444909052925
- Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**: 352–367. doi:10.1107/S0907444912001308
- Aravind L, de Souza RF, Iyer LM. 2010. Predicted class-I aminoacyl tRNA synthetase-like proteins in non-ribosomal peptide synthesis. *Biol Direct* **5**: 48. doi:10.1186/1745-6150-5-48
- Belin P, Moutiez M, Lautru S, Seguin J, Pernodet JL, Gondry M. 2012. The nonribosomal synthesis of diketopiperazines in tRNA-dependent cyclodipeptide synthase pathways. *Nat Prod Rep* **29**: 961–979. doi:10.1039/c2np20010d
- Bonnefond L, Arai T, Sakaguchi Y, Suzuki T, Ishitani R, Nureki O. 2011. Structural basis for nonribosomal peptide synthesis by an aminoacyl-tRNA synthetase paralog. *Proc Natl Acad Sci* **108**: 3912–3917. doi:10.1073/pnas.1019480108
- Borgman P, Lopez RD, Lane AL. 2019. The expanding spectrum of diketopiperazine natural product biosynthetic pathways containing cyclodipeptide synthases. *Org Biomol Chem* **17**: 2305–2314. doi:10.1039/C8OB03063D
- Borthwick AD. 2012. 2,5-Diketopiperazines: synthesis, reactions, medicinal chemistry, and bioactive natural products. *Chem Rev* **112**: 3641–3716. doi:10.1021/cr200398y
- Bourgeois G, Seguin J, Babin M, Belin P, Moutiez M, Mechulam Y, Gondry M, Schmitt E. 2018. Structural basis for partition of the cyclodipeptide synthases into two subfamilies. *J Struct Biol* **203**: 17–26. doi:10.1016/j.jsb.2018.03.001
- Byrne RT, Konevega AL, Rodnina MV, Antson AA. 2010. The crystal structure of unmodified tRNA^{Phe} from *Escherichia coli*. *Nucleic Acids Res* **38**: 4154–4162. doi:10.1093/nar/gkq133
- Byrne RT, Jenkins HT, Peters DT, Whelan F, Stowell J, Aziz N, Kasatsky P, Rodnina MV, Koonin EV, Konevega AL, et al. 2015. Major reorientation of tRNA substrates defines specificity of dihydrouridine synthases. *Proc Natl Acad Sci* **112**: 6033. doi:10.1073/pnas.1500161112
- Canu N, Moutiez M, Belin P, Gondry M. 2019. Cyclodipeptide synthases: a promising biotechnological tool for the synthesis of diverse 2,5-diketopiperazines. *Nat Prod Rep* **37**: 312–321. doi:10.1039/C9NP00036D
- Carter CW Jr, Wills PR. 2018. Hierarchical groove discrimination by Class I and II aminoacyl-tRNA synthetases reveals a palimpsest of the operational RNA code in the tRNA acceptor-stem bases. *Nucleic Acids Res* **46**: 9667–9683. doi:10.1093/nar/gky600
- Collaborative Computational Project Number 4. 1994. The CCP4 suite: programs from protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**: 760–763. doi:10.1107/S0907444994003112
- David G, Perez J. 2009. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J Appl Crystallogr* **42**: 892–900. doi:10.1107/S0021889809029288
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501. doi:10.1107/S0907444910007493
- Fonvielle M, de La Sierra-Gallay IL, El-Sagheer AH, Lecerf M, Patin D, Mellal D, Mayer C, Blanot D, Gale N, Brown T, et al. 2013. The structure of FemX_{WV} in complex with a peptidyl-RNA conjugate: mechanism of aminoacyl transfer from Ala-tRNA^{Ala} to peptidoglycan precursors. *Angew Chem Int Ed Engl* **52**: 7278–7281. doi:10.1002/anie.201301411
- Giegé R, Springer M. 2012. Aminoacyl-tRNA synthetases in the bacterial world. *EcoSal Plus* **7** doi:10.1128/ecosalplus.ESP-0002-2016
- Giessen TW, Marahiel MA. 2014. The tRNA-dependent biosynthesis of modified cyclic dipeptides. *Int J Mol Sci* **15**: 14610–14631. doi:10.3390/ijms150814610
- Gondry M, Sauguet L, Belin P, Thai R, Amouroux R, Tellier C, Tuphile K, Jacquet M, Braud S, Courcon M, et al. 2009. Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. *Nat Chem Biol* **5**: 414–420. doi:10.1038/nchembio.175

- Gouet P, Courcelle E, Stuart DI, Metoz F. 1999. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305–308. doi:10.1093/bioinformatics/15.4.305
- Guillon JM, Mechulam Y, Blanquet S, Fayat G. 1993. Importance of formylability and anticodon stem sequence to give a tRNA^{Met} an initiator identity in *Escherichia coli*. *J Bacteriol* **175**: 4507–4514. doi:10.1128/JB.175.14.4507-4514.1993
- Hajizadeh NR, Franke D, Jeffries CM, Svergun DI. 2018. Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data. *Sci Rep* **8**: 7204. doi:10.1038/s41598-018-25355-2
- Hol WG, van Duijnen PT, Berendsen HJ. 1978. The α -helix dipole and the properties of proteins. *Nature* **273**: 443–446. doi:10.1038/273443a0
- IUBMB. 1996. Prokaryotic and eukaryotic translation factors. Ad Hoc Nomenclature Subcommittee Report. *Biochimie* **78**: 1119–1122. doi:10.1016/S0300-9084(97)86738-7
- Jacques IB, Moutiez M, Witwinowski J, Darbon E, Martel C, Seguin J, Favry E, Thai R, Lecoq A, Dubois S, et al. 2015. Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity. *Nat Chem Biol* **11**: 721–727. doi:10.1038/nchembio.1868
- Juhling F, Morl M, Hartmann RK, Sprinzl M, Stadler PF, Putz J. 2009. tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res* **37**: D159–D162. doi:10.1093/nar/gkn772
- Kabsch WJ. 1988. Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J Appl Crystallogr* **21**: 916–924. doi:10.1107/S0021889888007903
- Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336**: 1030–1033. doi:10.1126/science.1218231
- Kobayashi T, Nureki O, Ishitani R, Yaremchuk A, Tukalo M, Cusack S, Sakamoto K, Yokoyama S. 2003. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat Struct Biol* **10**: 425–432. doi:10.1038/nsb934
- Konarev PV, Volkov VV, Petoukhov MV, Svergun DI. 2006. ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Crystallogr* **39**: 277–286. doi:10.1107/S0021889806004699
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40**: 658–674. doi:10.1107/S0021889807021206
- Mechulam Y, Blanquet S, Fayat G. 1987. Dual-level control of the *E. coli* pheST-himA operon expression: tRNA^{Phe}-dependent attenuation and transcriptional operator repressor control. *J Mol Biol* **197**: 453–470. doi:10.1016/0022-2836(87)90558-4
- Mechulam Y, Guillon L, Yatime L, Blanquet S, Schmitt E. 2007. Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. *Methods Enzymol* **430**: 265–281. doi:10.1016/S0076-6879(07)30011-6
- Meinzel T, Mechulam Y, Fayat G, Blanquet S. 1992. Involvement of the size and sequence of the anticodon loop in tRNA recognition by mammalian and *E. coli* methionyl-tRNA synthetases. *Nucleic Acids Res* **20**: 4741–4746. doi:10.1093/nar/20.18.4741
- Moutiez M, Schmitt E, Seguin J, Thai R, Favry E, Belin P, Mechulam Y, Gondry M. 2014a. Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. *Nat Commun* **5**: 5141. doi:10.1038/ncomms6141
- Moutiez M, Seguin J, Fonvielle M, Belin P, Jacques IB, Favry E, Arthur M, Gondry M. 2014b. Specificity determinants for the two tRNA substrates of the cyclodipeptide synthase AlbC from *Streptomyces noursei*. *Nucleic Acids Res* **42**: 7247–7258. doi:10.1093/nar/gku348
- Nissen P, Kjeldgaard M, Thirup S, Polekhina G, Reshetnikova L, Clark BFC, Nyborg J. 1995. Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* **270**: 1464–1472. doi:10.1126/science.270.5241.1464
- Piiaodov V, de Araújo E A, Neto M O, Craievich AF, Polikarpov I. 2019. SAXSMoW 2.0: online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. *Protein Sci* **28**: 454–463. doi:10.1002/pro.3528
- Rould MA, Perona JJ, Söll D, Steitz TA. 1989. Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA^{Gln} and ATP at 2.8 Å resolution. *Science* **246**: 1135–1142. doi:10.1126/science.2479982
- Sauguet L, Moutiez M, Li Y, Belin P, Seguin J, Le Du M-H, Thai R, Masson C, Fonvielle M, Pernodet J-L, et al. 2011. Cyclodipeptide synthases, a family of class-I aminoacyl-tRNA synthetase-like enzymes involved in non-ribosomal peptide synthesis. *Nucleic Acids Res* **39**: 4475–4489. doi:10.1093/nar/gkr027
- Schmitt E, Panvert M, Blanquet S, Mechulam Y. 1998. Crystal structure of methionyl-tRNA^{Met} transformylase complexed with the initiator formyl-methionyl-tRNA^{Met}. *EMBO J* **17**: 6819–6826. doi:10.1093/emboj/17.23.6819
- Schmitt E, Bourgeois G, Gondry M, Aleksandrov A. 2018. Cyclization reaction catalyzed by cyclodipeptide synthases relies on a conserved tyrosine residue. *Sci Rep* **8**: 7031. doi:10.1038/s41598-018-25479-5
- Schrodinger L. 2017. The PyMOL Molecular Graphics System, Version 2.0.0.
- Shen N, Guo L, Yang B, Jin Y, Ding J. 2006. Structure of human tryptophanyl-tRNA synthetase in complex with tRNA^{Trp} reveals the molecular basis of tRNA recognition and specificity. *Nucleic Acids Res* **34**: 3246–3258. doi:10.1093/nar/gkl441
- Storoni LC, McCoy AJ, Read RJ. 2004. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **60**: 432–438. doi:10.1107/S0907444903028956
- Suto K, Shimizu Y, Watanabe K, Ueda T, Fukai S, Nureki O, Tomita K. 2006. Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J* **25**: 5942–5950. doi:10.1038/sj.emboj.7601433
- Svergun DI, Barberato C, Koch MHJ. 1995. CRYSOLE – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**: 768–773. doi:10.1107/S0021889895007047
- Vetting MW, Hegde SS, Blanchard JS. 2010. The structure and mechanism of the Mycobacterium tuberculosis cyclodityrosine synthetase. *Nat Chem Biol* **6**: 797–799. doi:10.1038/nchembio.440
- Vonrhein C, Tickle IJ, Flensburg C, Keller P, Paciorek W, Sharff A, Bricogne G. 2018. Advances in automated data analysis and processing within autoPROC, combined with improved characterisation, mitigation and visualisation of the anisotropy of diffraction limits using STARANISO. *Acta Crystallogr* **74**: a360. doi:10.1107/S010876731809640X
- Yang XL, Otero FJ, Ewalt KL, Liu J, Swairjo MA, Kohrer C, RajBhandary UL, Skene RJ, McRee DE, Schimmel P. 2006. Two conformations of a crystalline human tRNA synthetase-tRNA complex: implications for protein synthesis. *EMBO J* **25**: 2919–2929. doi:10.1038/sj.emboj.7601154
- Yaremchuk A, Kriklyvi I, Tukalo M, Cusack S. 2002. Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J* **21**: 3829–3840. doi:10.1093/emboj/cdf373
- Zubay G. 1962. The isolation and fractionation of soluble ribonucleic acids. *J Mol Biol* **4**: 347–356. doi:10.1016/S0022-2836(62)80015-1



RNA

A PUBLICATION OF THE RNA SOCIETY

Structural basis of the interaction between cyclodipeptide synthases and aminoacylated tRNA substrates

Gabrielle Bourgeois, Jérôme Seguin, Morgan Babin, et al.

RNA 2020 26: 1589-1602 originally published online July 17, 2020
Access the most recent version at doi:[10.1261/rna.075184.120](https://doi.org/10.1261/rna.075184.120)

Supplemental Material <http://rnajournal.cshlp.org/content/suppl/2020/07/17/rna.075184.120.DC1>

References This article cites 51 articles, 8 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/26/11/1589.full.html#ref-list-1>

Open Access Freely available online through the *RNA* Open Access option.

Creative Commons License This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

SMART® cDNA + library prep:
now all from **one source**



To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>
