SUPPORTING INFORMATION

FOR

Cyclodipeptide Synthases of the NYH Subfamily Recognize tRNA Using an α-Helix Enriched with Positive Residues

Anastasia Croitoru[1], Morgan Babin[2], Hannu Myllykallio[1], Muriel Gondry[2], and Alexey Aleksandrov[1*]

[1]Laboratoire d'Optique et Biosciences (CNRS UMR7645, INSERM U1182), Ecole Polytechnique, Institut polytechnique de Paris, F-91128 Palaiseau, France

[2]Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198 Gif-sur-Yvette cedex, France

*Email: Alexey.Aleksandrov@polytechnique.edu

METHODS AND MATERIALS

Homology modeling of the AlbC:tRNA complex based on the TyrRS-tRNA structure

Available three structures of the TyrRS-tRNA complexes were retrieved from the Protein Data Bank with PDB reference codes 1H3E (1), 2DLC (2) and 1J1U (3). Although AlbC has only a low sequence identity with TyrRSs, below 16%, it has a significant structural resemblance with the RMSD of backbone Cα atoms varying from 3.4 to 3.9 Å (Table S6). The structure of *Thermus thermophilus* TyrRS:tRNA (it will be referred as TyrRS in what follows) with PDB reference code 1H3E was chosen for modelling of the AlbC:tRNA complex as the most complete crystal structure and having a high sequence similarity with the AlbC protein.

TyrRS is a functional homodimer with cross-subunit tRNA binding. It recognizes the cognate Tyr-tRNA anticodon and variable loop through its C-terminal domain and does not require a distortion of the 3'-extremity of the tRNA from its helical path to enter the active site (1). The AlbC:Phe-tRNA^{Phe} complex was obtained by superposing AlbC catalytic domain on the TyrRS equivalent domain and Phe-tRNA^{Phe} on tRNA^{Tyr} unit respectively. The structures for AlbC, Phe-tRNA^{Phe} and TyrRS:tRNA^{Tyr} complex were obtained from PDB with codes 4Q24 (4), 4YCO chain D (5) and 1H3E (1), respectively. To remove any clashes between Phe-tRNA^{Phe} and AlbC, Phe-tRNA^{Phe} was manually translated to obtain a non-overlapping system. In the AlbC:Phe-tRNA^{Phe} model the aminoacylated 3' of tRNA was placed in the catalytic site of CDPS in a similar orientation to the 3' position inside the catalytic site of TyrRS. The system was prepared for MD simulations with the same protocol described in the main text for docking models. Finally, the Phe-tRNA^{Phe} orientation was relaxed with targeted molecular dynamics simulations.

Targeted molecular dynamics (TMD) simulations were performed for one-ns to relax possible clashes in the modelled AlbC:tRNA complex that may result from the inaccurate positioning of the protein. Steering forces with a force constant k of 10 kcal·mol^{-1}·Å^{-2} were applied on the heavy atoms in TMD. At each time-step, the Root Mean Square (RMS) distance between MD coordinates and the target structure is computed (after first aligning the target structure to the current coordinates). The force on each atom is given by the gradient of the potential U:

$$U = \frac{k}{2N}[RMS(t) - RMS^*(t)]^2, \qquad \text{[Eq 1]}$$

where $RMS(t)$ is the best-fit RMS deviation of the current coordinates from the coordinates of the targeted structure, and $RMS(t)^*$ evolves linearly from the initial RMSD at the first step to the final RMSD at the last step. N is the number of targeted atoms; k is the force constant. Forces due to TMD were applied on heavy atoms of AlbC within 10 Å of residues important for tRNA interaction in TyrRS. Residues of TyrRS important for its interaction with tRNA^{Tyr} were identified via the component analysis. For the targeted structure for the complex, the structure of AlbC:Phe-tRNA^{Phe} produced by

superimposing AlbC and tRNA on the TyrRS:$tRNA^{Tyr}$ structure was used. The electrostatic contribution to the protein:tRNA binding free energy was computed by the PB/SA approach (see Material and Methods).

TyrRS residues within 6 Å of tRNA used for this analysis indicate the largest contributions to the binding free energy (Table S7) made by Arg205, Arg209, Lys159, Arg198, Arg155, Asn151. In particular, Arg209 with -10.23 kcal·mol^{-1} and Arg205 with -19.02 kcal·mol^{-1} prove to have the principal role in tRNA binding. These residues belong to the clusters 148-154 and 198-211 identified as important for tRNA recognition (1). Positive values observed for E154, E156 and E206 indicated that these residues are an impediment for the tRNA interaction. Being negatively charged they repulse the negatively charged nucleic acids. However, this repulsion is compensated by the presence of positive and polar protein residues. In AlbC residues homologues of TyrRS residues important for tRNA binding are Arg160, Arg154, His203, Arg214, Arg215 and Arg220. They served to orientate the AlbC-tRNA in the TMD simulations.

Poisson-Boltzmann (PB) binding free energies for ZDOCK docking structures

In addition to the ZDOCK score, the Poisson-Boltzmann binding free energies were estimated for 2,000 docking structures. For these calculations, the structures of the complex were relaxed to remove possible clashes between atoms and to allow protein sidechains to adjust and create interactions with tRNA. First, all atoms of the complex were energetically minimized for individual structures using 2,000 steps of steepest-descent minimization in CHARMM. The protein and tRNA backbone heavy atoms were restrained to the initial structure using a 1.0-kcal·mol^{-1}·$Å^{-2}$ force constant. Starting from the minimized structures Molecular Dynamics simulations were performed for 200-ps in NAMD with the restraints applied to the tRNA heavy atoms with a force constant of 1.0-kcal·mol^{-1}·$Å^{-2}$. A cutoff distance of 15 Å for electrostatic and van der Waals calculations with the energy switching between 13 Å and 15 Å was used; the dielectric constant of 4 was used to introduce the solvent screening. The Poisson-Boltzmann binding free energy calculations were performed on the structures after the MD simulations with the CHARMM program and using the same parameters and setup described in the Method section of the main text. To evaluate contribution of the α4 helix to the binding, the PB binding calculations were performed on structures excluding the sidechains of the α4 residues.

Force field determination for aminoacyl group of tRNA

As a part of this study, we report a force field model for the aminoacyl group of tRNA. The developed force field is compatible with the CHARMM36 (C36) (6) and with the TIP3P water model (7–9). The parametrization procedure was similar to the work described previously (10). The protocol for the force field development is presented briefly below, more details can be found in ref. (10).

Parametrization of non-bonded terms

Consistent with the development of the CHARMM force field, atomic charges were derived targeting interactions between the model compound and individual water molecules, and the dipole

moment of the model compound as well as quantum mechanical (QM) electrostatic potential. The charge optimization was performed on the compound structures optimized with the B3LYP functional (11) and 6-31G(d) basis set (12). The Gaussian09 (Gaussian, Inc, Wallingford CT, 2009) program was used for all QM calculations. Atoms of the model compound that can participate in hydrogen bonds were probed by individual water molecules placed in idealized linear orientations (8). Each water-compound structure was then optimized by varying the interaction distance, to find the energy minimum for the water position. Calculations were done at the HF/6-31G(d) level (6, 8) as for the C36. In accord with the standard CHARMM parametrization protocol (8), the *ab initio* interaction energies and the minimum interaction distance were corrected by the empirical factors. The molecular dipole moment was included to provide additional target data for the optimization of the atomic charges of neutral compounds (10). The dipole moment was calculated in vacuum at the same B3LYP/6-31G(d) level.

Parametrization of bonded terms

Following our previous work, missing parameters for all bonded terms were optimized using potential energy scans (PES) (13). In particular, an adiabatic PES scan was performed for each degree of freedom that has adjustable parameters in the force field. The same protocol was applied to parametrize the terms associated with soft dihedrals in CHARMM. The bonded parameters were varied to improve the agreement between CHARMM and QM PES. All PES scans were performed at the B3LYP/6-31G(d) level.

Results

Parametrization of aminoacyl group Phe-tRNA

Water interactions with model compounds are given in Tables S8 and S9 for alanine amino acid with the methylated C-terminus and standard N-terminus in the deprotonated and protonated forms, respectively. Overall, good agreement was achieved for interactions with water molecules, the Root Mean Square Deviation between interaction energies computed with CHARMM and QM are 0.48 $kcal \cdot mol^{-1}$ and 0.64 $kcal \cdot mol^{-1}$ for the neutral and protonated forms respectively. In addition, the dipole moment was well reproduced for the neutral form. The dipole moment is 5.83 D and 5.80 D with the QM at the B3LYP/6-31G* level and CHARMM models respectively. Interaction distances are also in a good agreement with the QM results, the RMS deviation for interaction distances is 0.16 and 0.23 Å for the neutral and protonated forms respectively.

Figure S11 demonstrates the agreement for the PES scan for the important dihedral angles associated with the bond connecting the deprotonated on the N terminus alanyl group with the ribose group. The force field model reasonably reproduces the minima as well as the height of the barriers between the minima.

Tables

Table S1. Structurally characterized CDPSs.

Protein	Source organism	Ref.
AlbC	*Streptomyces noursei*	(4, 14)
Rv2275	*Mycobacterium tuberculosis*	(15)
YvmC	*Bacillus licheniformis*	(16)
Nbra-CDPS	*Nocardia brasiliensis*	(17)
Rgry-CDPS	*Rickettsiella grylli*	(17)
Fdum-CDPS	*Fluoribacter dumoffii*	(17)
Shae-CDPS	*Staphylococcus haemolyticus*	(17)

Table S2. RMS deviation (Å) computed for tRNA backbone between the proposed docked models. The RMS deviation was calculated for the tRNA backbone heavy atoms after the models were superimposed using the protein backbone atoms. Values lower than 20 Å are highlighted.

Model	TyrRS	1	2	3	4	5	6	7	8	9	10
TyrRS	-	85.2	74.7	77.5	82.5	13.4	87.4	84.0	83.7	16.4	81.9
1		-	27.0	25.9	17.7	89.0	13.3	20.4	14.2	84.6	11.3
2			-	28.7	16.0	77.9	29.3	25.3	18.8	70.9	21.6
3				-	26.4	80.6	21.3	17.2	23.6	75.0	19.1
4					-	84.8	16.7	15.6	5.3	79.1	17.9
5						-	90.0	86.0	86.3	11.4	86.1
6							-	10.3	11.9	85.4	15.8
7								-	11.9	80.6	17.7
8									-	80.8	14.4
9										-	80.8
10											-

Table S3. Poisson-Boltzmann/Surface Area binding free energies. The vibrational entropy was estimated by Normal Mode Analysis. Energies are given in kcal·mol^{-1} relative to model 7. The uncertainty of calculations is given in parentheses.

Model	PB/SA	Entropy	Total
1	5.3 (3.8)	8.4 (5.7)	13.7 (6.9)
2	6.2 (5.2)	12.4 (5.1)	18.6 (7.3)
3	12.5 (3.2)	10.8 (1.7)	23.3 (3.6)
4	6.8 (5)	7.7 (6.8)	13.9 (8.5)
5	25.7 (3)	2.1 (5.7)	27.8 (6.5)
6	10.8 (3.2)	13.7 (2.7)	24.5 (4.2)
7	0 (2.9)	0.0 (1.6)	0.0 (3.3)
8	4.4 (3.3)	14.8 (4.0)	19.2 (5.2)
9	16 (3.1)	0.4 (3.7)	16.4 (4.8)
10	11.3 (3.1)	7.5 (6.9)	18.8 (7.6)

Table S4. Contributions of individual residues to the binding free energy in kcal·mol^{-1} Component analysis was performed for residues within 6 Å from the tRNA. The energies were averaged over ten snapshots taken every 20 ns after 100 ns of equilibration for models 2 and 7, and each 10 ns after 100 ns of equilibration for model 9.

Residue	Model 2	Model 7	Model 9
K46	**-8.2** (0.9)	-1.2 (0.6)	-0.3 (0.2)
R91	-1.3 (0.5)	**-8.0** (3.9)	0.0 (<0.1)
K94	**-2.6** (1.1)	**-8.3** (2.7)	0.0 (<0.1)
D95	**2.3** (0.8)	**4.0** (0.9)	0.0 (<0.1)
R98	**-14.2** (6.4)	**-9.9** (2.4)	0.0 (<0.1)
R99	**-8.0** (2.7)	**-6.7** (3.2)	-0.2 (0.1)
R102	**-7.3** (2.0)	**-8.9** (1.5)	0.0 (<0.1)
D182	**-6.7** (1.1)	**-4.7** (1.9)	**-5.8** (1.3)
R231	-0.9 (0.3)	-0.2 (0.1)	**-8.3** (0.8)

Absolute binding free energy contributions of AlbC residues higher than 2 kcal·mol^{-1} are given in bold.

Table S5. *In vivo* production of cFL and cFF by AlbC and its variants.

AlbC and its variants	Cyclodipeptides produced	Area (214 nm)*		Relative activity compared to the wild-type (%)
		E 1	E2	
Wild-type	cFL	7043	7174	-
	cFF	3229	3022	-
K46A	cFL	6535	7551	100 ± 8
	cFF	3810	3705	120 ± 2
D95A	cFL	2325	3173	39 ± 6
	cFF	1432	1297	44 ± 1

The cyclodipeptide-synthesizing activities were determined by LC-MS analyses of bacterial culture supernatants. *Peak areas at 214 nm correspond to two-independent experiments, E1 and E2. cFL and cFF were eluted at retention times of 23.4 and 24.7 min, respectively. The precursor ions were respectively at m/z 261 and 295 for cFL and cFF.

Table S6. Crystal structures of Tyrosyl-tRNA synthetase in complex with tRNA compared to AlbC. The RMS deviation was computed between backbone Cα's after structural superposition on the crystal structure PDB: 4Q24.

PDB	Organism	RMSD, Å	Sequence identity
1H3E (1)	*Thermus thermophilus*	3.9	16%
2DLC (1)	*Saccharomyces cerevisiae*	3.6	15%
1J1U (3)	*Methanococcus jannaschii*	3.4	10%

Table S7. Contribution of individual residues in kcal·mol^{-1} to the binding free energy of TyrRS to tRNA computed using the PB/SA model

TyrRS residues	Q151	E154	R155	E156	K159	K160	R 198	R205	E206	R209
AlbC homologues	E150	V153	R160	Q155	R154	P161	H203	R214	E216	R220/R215
Free energy	-2.41	5.48	-3.44	12.25	-5.68	-3.42	-5.65	-19.02	7.47	-10.23

Table S8. Interactions between a probe water and selected sites of alanine amino acid with the methylated C-terminus and standard N-terminus in the neutral form

	ab initio/force field results				*ab initio*/force field results		
probe site	energy (kcal · mol^{-1})	distance (Å)	angle (º)	probe site	energy (kcal · mol^{-1})	distance (Å)	angle (º)
HN1-N	-3.45/-3.76	2.11/1.95	0.0	HB2-CB	-1.87/-1.32	2.48/2.69	90.0
HN1-N	-3.58/-3.79	2.09/1.95	45.0	HB3-CB	-1.02/-1.02	2.48/2.67	0.0
HN1-N	-3.30/-3.70	2.11/1.95	90.0	HB3-CB	-0.77/-0.90	2.61/2.70	90.0
HN1-N	-3.18/-3.68	2.13/1.9	135.0	OT1-C	-4.82/-5.11	1.89/1.79	0.0
HC1-CT	-1.17/-0.93	2.39/2.63	0.0	OT1-C	-5.25/-5.57	1.87/1.78	90.0
HC1-CT	-1.37/-0.96	2.32/2.62	90.0	OT1-C	-4.76/-5.01	1.89/1.79	180.0
HC2-CT	-1.80/-1.23	2.40/2.66	0.0	OT1-C	-4.72/-4.93	1.89/1.79	270.0
HC2-CT	-1.75/-1.21	2.41/2.67	90.0	N-HN1	-6.80/-6.21	1.90/1.91	0.0
HC3-CT	-1.01/-0.95	2.45/2.65	0.0	N-HN1	-7.29/-6.40	1.89/1.91	90.0
HC3-CT	-1.43/-1.13	2.36/2.63	90.0	N-HN1	-6.65/-5.81	1.90/1.91	180.0
HA-CA	-0.99/-0.75	2.45/2.63	0.0	N-HN1	-6.03/-5.49	1.92/1.92	270.0
HA-CA	-0.78/-0.66	2.52/2.64	90.0	OT2-C	-3.37/-3.12	1.94/1.95	0.0
HB1-CB	-1.50/-1.22	2.41/2.64	0.0	OT2-C	-4.10/-3.24	1.90/1.96	90.0
HB1-CB	-1.07/-1.02	2.48/2.66	90.0	OT2-C	-2.79/-2.13	2.00/2.01	180.0
HB2-CB	-1.85/-1.36	2.50/2.69	0.0	OT2-C	-3.50/-2.48	1.92/1.97	270.0

Table S9. Interactions between a probe water and selected sites of alanine amino acid with the methylated C-terminus and standard N-terminus in the protonated form

	ab initio/force field results				*ab initio*/force field results		
probe site	energy (kcal · mol^{-1})	distance (Å)	angle (º)	probe site	energy (kcal · mol^{-1})	distance (Å)	angle (º)
HN1-N	-17.60/-17.05	1.72/1.74	0.0	HC1-CT	-5.51/-5.01	2.23/2.53	90.0
HN1-N	-17.78/-17.12	1.72/1.74	45.0	HC2-CT	-5.54/-4.91	2.23/2.53	0.0
HN1-N	-17.63/-17.06	1.72/1.74	90.0	HC2-CT	-5.65/-4.97	2.22/2.53	90.0
HN1-N	-17.45/-16.99	1.72/1.75	135.0	HC3-CT	-4.77/-4.65	2.30/2.55	0.0
HN2-N	-16.96/-17.01	1.75/1.76	0.0	HC3-CT	-5.50/-5.02	2.23/2.53	90.0
HN2-N	-16.79/-16.88	1.75/1.76	45.0	HB1-CB	-7.69/-7.11	2.23/2.50	0.0
HN2-N	-17.12/-17.02	1.74/1.76	90.0	HB1-CB	-7.62/-7.06	2.23/2.50	90.0
HN2-N	-17.27/-17.15	1.74/1.76	135.0	HB2-CB	-7.31/-6.68	2.26/2.53	0.0
HN3-N	-3.17/-3.54	3.53/3.62	0.0	HB2-CB	-7.64/-6.84	2.23/2.52	90.0
HN3-N	-3.07/-3.51	3.62/3.65	45.0	HB3-CB	-6.08/-5.55	2.21/2.55	0.0
HN3-N	-3.16/-3.51	3.49/3.61	90.0	HB3-CB	-5.97/-5.57	2.23/2.55	90.0
HN3-N	-3.31/-3.54	3.37/3.58	135.0	HA-CA	-10.02/-8.49	2.05/2.42	0.0
HC1-CT	-4.81/-4.66	2.29/2.55	0.0	HA-CA	-10.41/-8.65	2.03/2.42	90.0

Figures

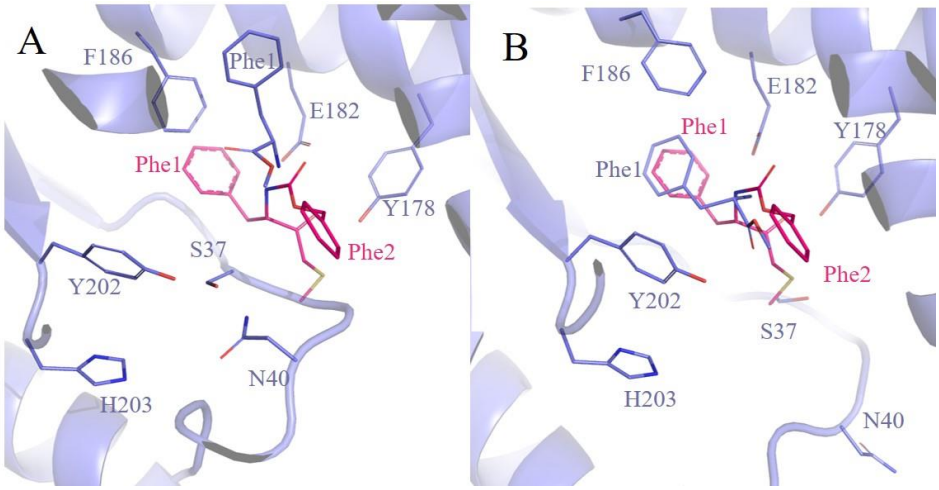

Figure S1. The orientation of the tRNA aminoacyl group in model 7 (A) and 9 (B) at the end of 30-ns MD simulations. The position of the dipeptide analogue in the experimental structure (PDB: 4Q24) after it was superimposed using the protein backbone heavy atoms is shown in pink.

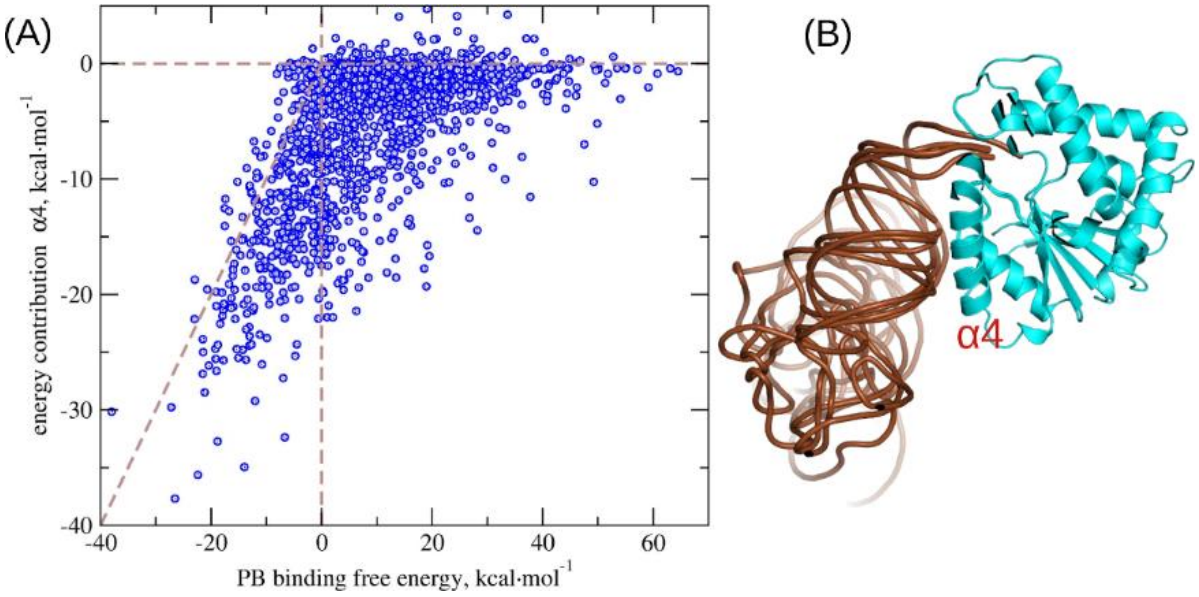

Figure S2. Poisson-Boltzmann (PB) binding free energies of 2,000 ZDOCK models for the AlbC:tRNA complex (A); five docked structures with the lowest PB energies (B). The binding free energies are plotted against the binding contribution of the α4 helix to the total binding free energy; the diagonal dashed line shows the linear correlation between the α4 contribution and the total binding energy. The five docking structures with the lowest PB energies on panel B were selected among the solution with the position of the tRNA aminoacyl group compatible with the experimental structure with the dipeptide analogue (PDB: 4Q24).

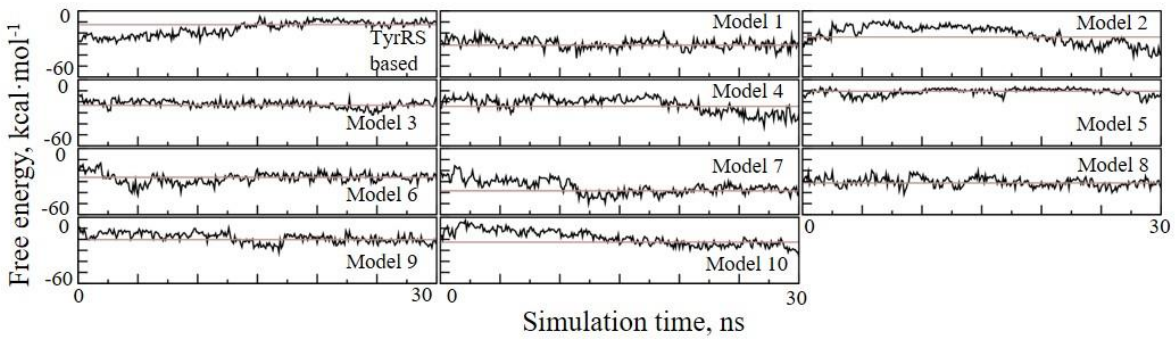

Figure S3. AlbC:tRNA binding free energy. The binding free energy was computed every 0.1 ns in MD simulations of the model built from AlbC-TyrRS homology and ten docking models. Average binding free energy between 10 ns and the end of MD simulations is represented by the gray line.

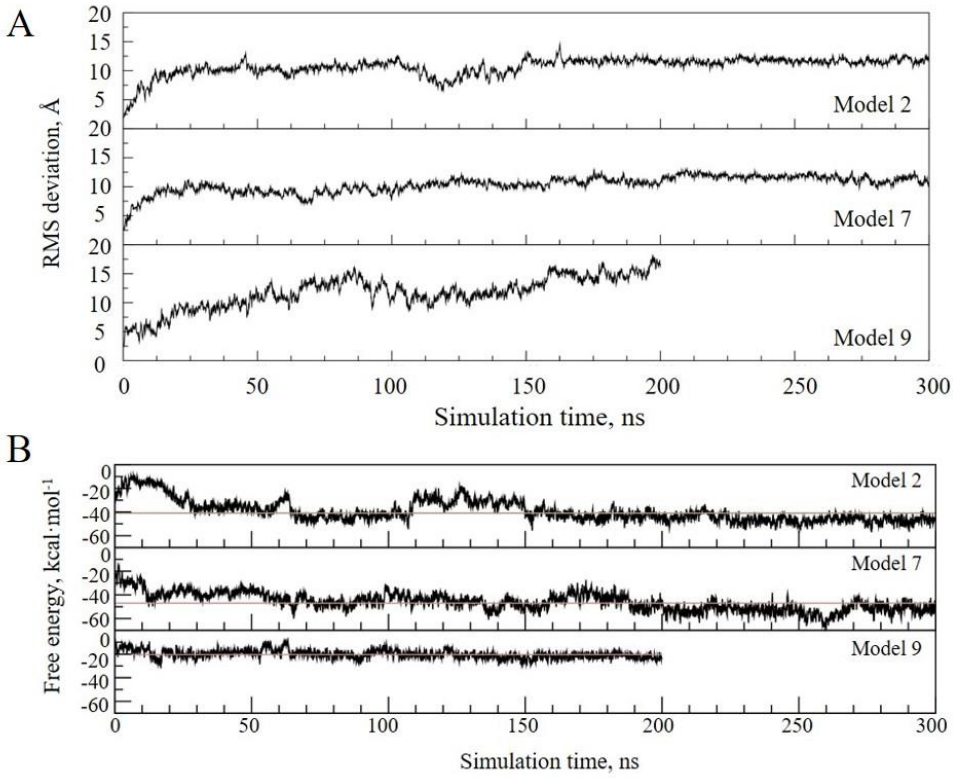

Figure S4. RMS deviation and binding free energy for models 7, 8 and 9. A) AlbC:tRNA RMS deviation. The RMS deviation determined every 0.1 ns of the model containing the complete tRNA built from models 7 and 9 is shown. B) AlbC:tRNA binding free energy. The free binding energy determined every 0.1 ns by PBSA calculations are shown for three docking models. Average binding free energy between 20 ns and the end of MD simulations is represented by a brown line.

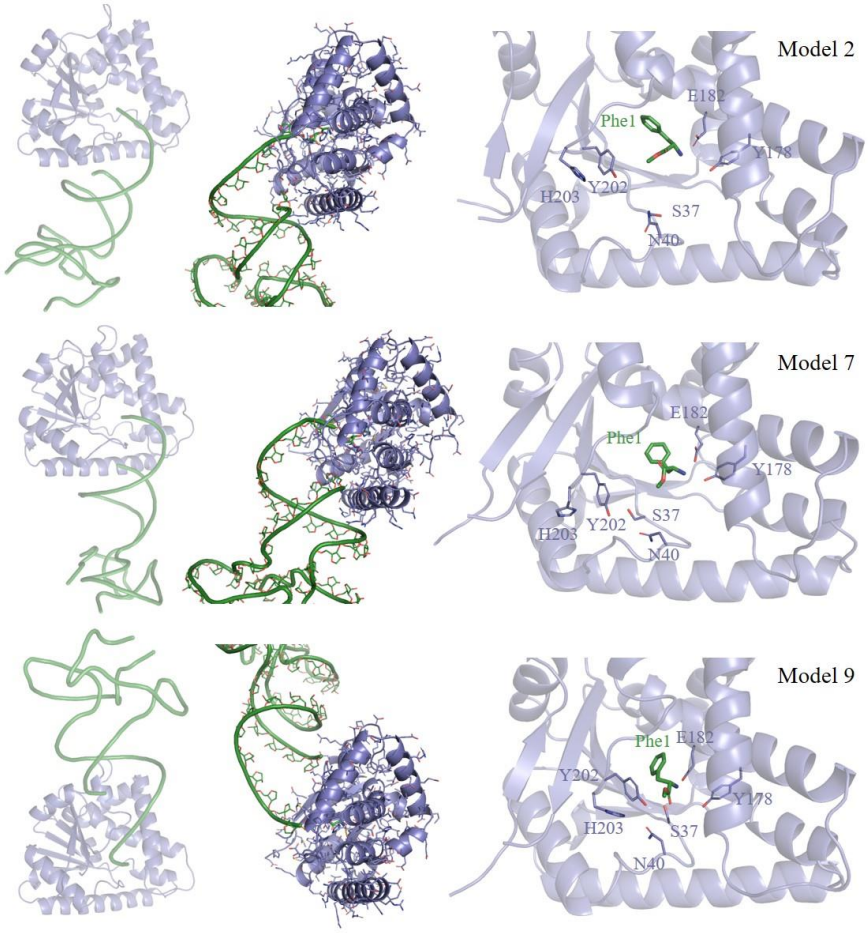

Figure S5. The models of the AlbC:tRNA complex with the lowest binding free energy. The left and middle panels show an overview of the complex, while the right panels show a closeup view of active site. The protein is in blue and tRNA is shown in green. In each panel AlbC has the same orientation.

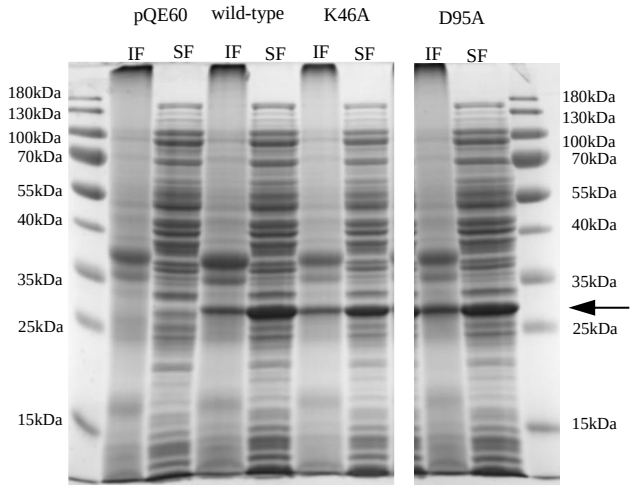

Figure S6. Analysis of the expression of AlbC wild-type, K46A and D95Avariants. Soluble (SF) and insoluble fractions (IF) were analyzed by 12% SDS-PAGE with Coomassie blue staining. pQE60 corresponds to the empty vector. The arrow indicates the wild-type and mutated AlbC expressed in soluble and insoluble fractions. Two independent experiments were analyzed and gave the same results.

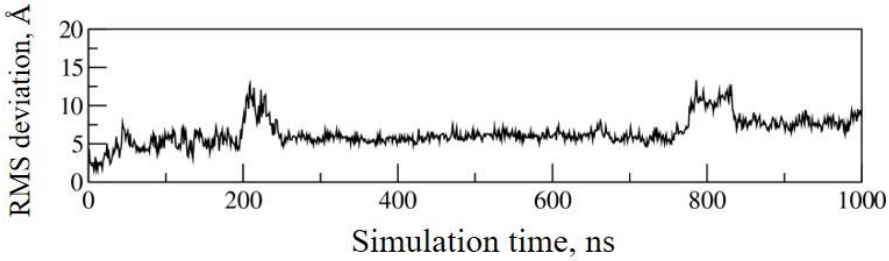

Figure S7. RMS deviation for AlbC:tRNA backbone atoms observed in MD simulations of the complete model. The RMS deviation determined every 0.1 ns of the model containing the complete tRNA is shown.

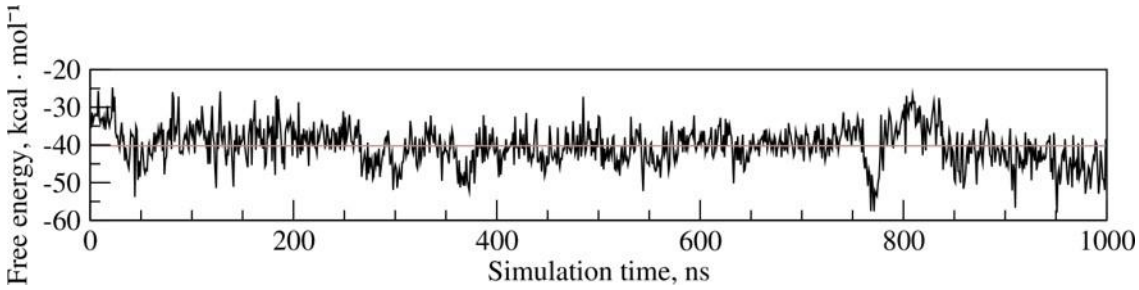

Figure S8. AlbC:tRNA binding free energy observed in the 1-μs MD simulation. The binding free energy was computed every 1 ns for the final model. The average value for the binding free energy is shown by the grey line.

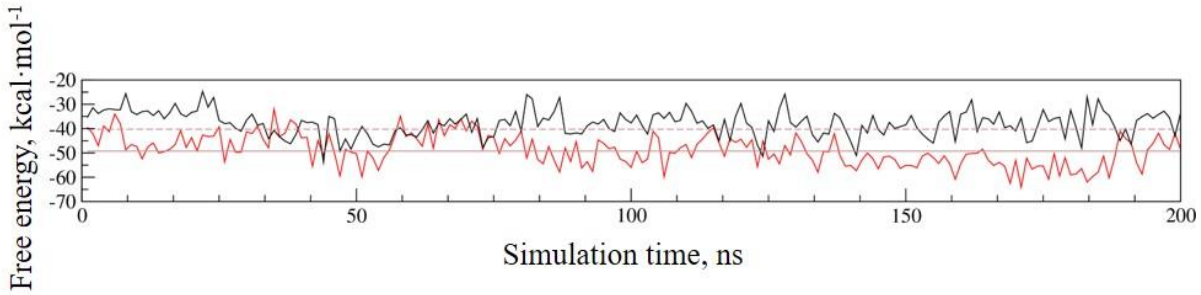

Figure S9. AlbC:tRNA binding free energy for the mutant D95A. The binding free energy was calculated every one-ns; the binding free energy for Asp95Ala variant is shown in red and for wild type is shown in black. Average binding free energy between 20 ns and the end of MD simulations is represented by a dashed grey line for wild type and a continuous grey line for D95A mutant.

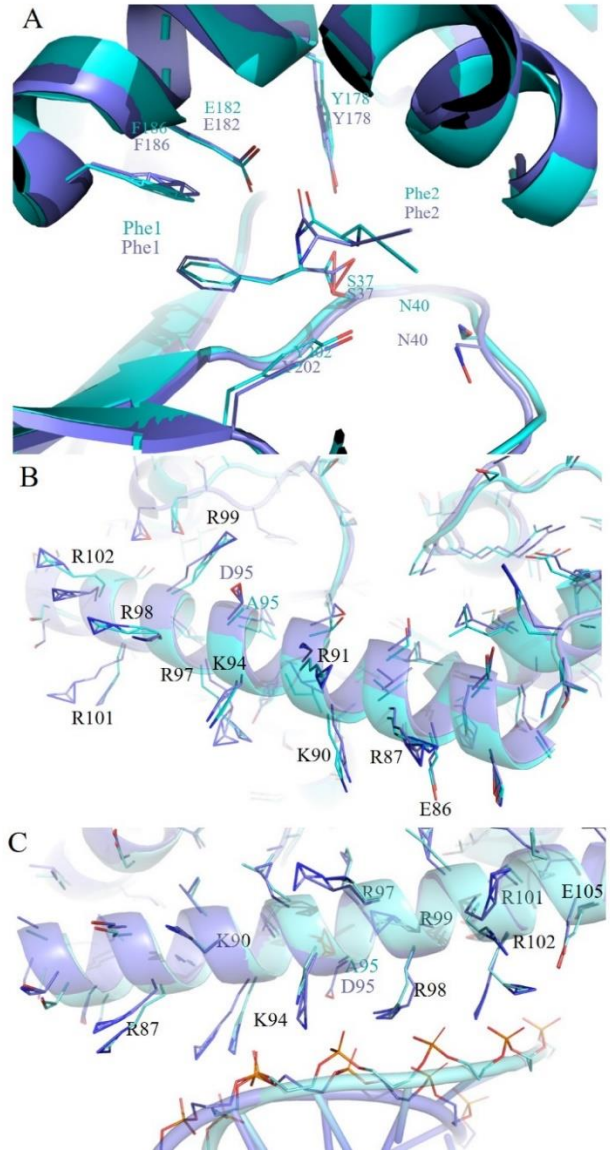

Figure S10. Average structures observed in MD simulations with the wild-type (WT) protein and the Asp95Ala variant. The WT protein is shown in blue and Asp95Ala in cyan in all panels. A) Close up view of the binding pocket structure and B) surface residues observed in MD simulations with AlbC in the dipeptide intermediate state without tRNA; C) the structure of the α4:tRNA interface in MD simulations of the WT and the Asp95Ala variant in complex with tRNA. In all cases the Asp95Ala and WT structures are very similar.

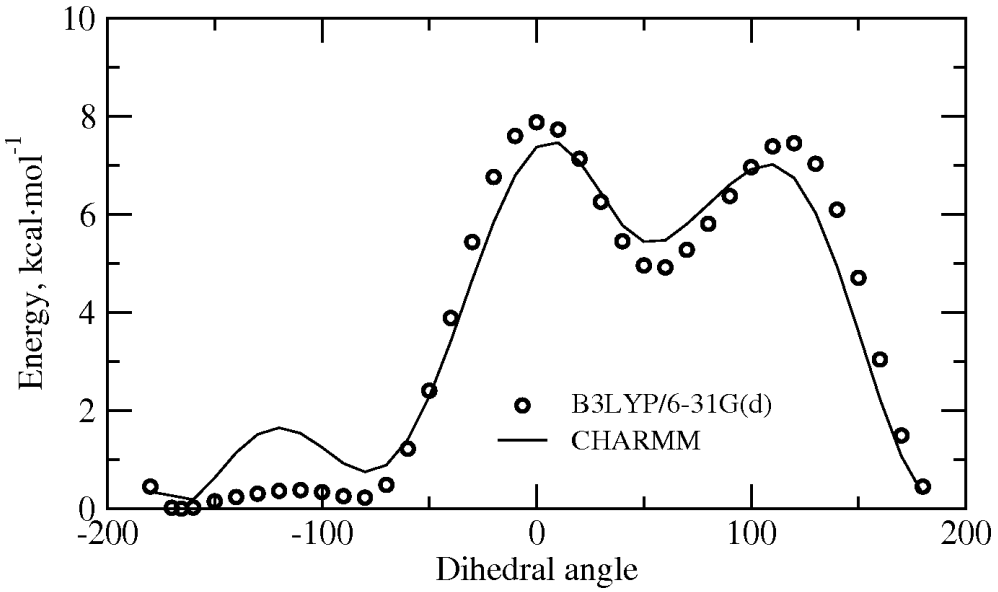

Figure S11. Potential energy scans along dihedral angles connecting the deprotonated on the N terminus alanyl group with the ribose group.

References

1. Yaremchuk, A. (2002) Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J.* **21**, 3829–3840
2. Tsunoda, M., Kusakabe, Y., Tanaka, N., Ohno, S., Nakamura, M., Senda, T., Moriguchi, T., Asai, N., Sekine, M., Yokogawa, T., Nishikawa, K., and Nakamura, K. T. (2007) Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms. *Nucleic Acids Res.* **35**, 4289–4300
3. Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Mol. Biol.* **10**, 425–432
4. Moutiez, M., Schmitt, E., Seguin, J., Thai, R., Favry, E., Belin, P., Mechulam, Y., and Gondry, M. (2014) Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. *Nat. Commun.* **5**, 5141
5. Byrne, R. T., Konevega, A. L., Rodnina, M. V., and Antson, A. A. (2010) The crystal structure of unmodified tRNA Phe from Escherichia coli. *Nucleic Acids Res.* **38**, 4154–4162
6. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell, A. D. (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690
7. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935
8. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B.* **102**, 3586–3616

9. Neria, E., Fischer, S., and Karplus, M. (1996) Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **105**, 1902–1921
10. Xu, Y., Vanommeslaeghe, K., Aleksandrov, A., MacKerell, A. D., and Nilsson, L. (2016) Additive CHARMM force field for naturally occurring modified ribonucleotides: CHARMM Potential Energy Function. *J. Comput. Chem.* **37**, 896–912
11. Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652
12. Krishnan, R., Binkley, J. S., Seeger, R., and Pople, J. A. (1980) Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **72**, 650–654
13. Aleksandrov, A. (2019) A Molecular Mechanics Model for Flavins. *J. Comput. Chem.* **40**, 2834–2842
14. Sauguet, L., Moutiez, M., Li, Y., Belin, P., Seguin, J., Le Du, M.-H., Thai, R., Masson, C., Fonvielle, M., Pernodet, J.-L., Charbonnier, J.-B., and Gondry, M. (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
15. Vetting, M. W., Hegde, S. S., and Blanchard, J. S. (2010) The structure and mechanism of the Mycobacterium tuberculosis cyclodityrosine synthetase. *Nat. Chem. Biol.* **6**, 797–799
16. Bonnefond, L., Arai, T., Sakaguchi, Y., Suzuki, T., Ishitani, R., and Nureki, O. (2011) Structural basis for nonribosomal peptide synthesis by an aminoacyl-tRNA synthetase paralog. *Proc. Natl. Acad. Sci.* **108**, 3912–3917
17. Bourgeois, G., Seguin, J., Babin, M., Belin, P., Moutiez, M., Mechulam, Y., Gondry, M., and Schmitt, E. (2018) Structural basis for partition of the cyclodipeptide synthases into two subfamilies. *J. Struct. Biol.* **203**, 17–26

The developed parameters are presented in the following section

```
* CHARMM36/C toppar stream file for aminoacyl group of aminoacyl-tRNA
* june/2018, Alexey Aleksandrov
*

read rtf card append
*  force field model for aminoacyl group of aminoacyl-tRNA, deprotonated form
*
36 1

PRES ADD1              0.000 ! patch to create aminoacyl group of tRNA
GROUP
ATOM C2'     CN7B      0.147
ATOM H2''    HN7       0.090
ATOM O2'     OS       -0.399
ATOM N       NH2      -0.891
ATOM HN1     H         0.349
ATOM HN2     H         0.349
ATOM CA      CT1       0.078
ATOM HA      HB1       0.090
ATOM C       CD        0.677
ATOM O       OB       -0.490
GROUP
ATOM CB      CT3      -0.270
ATOM HB1     HA3       0.090
ATOM HB2     HA3       0.090
ATOM HB3     HA3       0.090
DELE ATOM H2'
BOND      O2'       C
BOND        N     HN1       N     HN2       N      CA
BOND       CA      HA      CA       C      CA      CB
BOND        C       O
BOND       CB     HB1      CB     HB2      CB     HB3
IMPR C CA O2' O

PRES ADDY              1.000 ! phenylalanyl group of phe-tRNA
GROUP
ATOM C2'     CN7B      0.256
ATOM H2''    HN7       0.090
ATOM O2'     OS       -0.321
ATOM N       NH2      -0.502
ATOM HN1     H         0.338
ATOM HN2     H         0.338
ATOM HN3     H         0.338
ATOM CA      CT1       0.435
ATOM HA      HB1       0.008
ATOM C       CD        0.508
ATOM O       OB       -0.488
GROUP
ATOM CB      CT2      -0.18
ATOM HB1     HA2       0.09   !         HD1  HE1
ATOM HB2     HA2       0.09   !          |    |
GROUP                         !    HB1  CD1--CE1
ATOM CG      CA        0.00   !    |    //     \\
GROUP                         !  --CB--CG      CZ--HZ
ATOM CD1     CA       -0.115  !    |    \  __  /
ATOM HD1     HP        0.115  !    HB2  CD2--CE2
GROUP                         !          |    |
ATOM CE1     CA       -0.115  !         HD2  HE2
ATOM HE1     HP        0.115
GROUP
ATOM CZ      CA       -0.115
ATOM HZ      HP        0.115
GROUP
ATOM CD2     CA       -0.115
ATOM HD2     HP        0.115
GROUP
ATOM CE2     CA       -0.115
ATOM HE2     HP        0.115

DELE ATOM H2'
BOND      O2'       C
BOND        N     HN1       N     HN2       N     HN3
BOND        N      CA
BOND       CA      HA      CA       C
BOND        C       O
IMPR C CA O2' O

BOND CB  CA   CG CB   CD2 CG   CE1 CD1
BOND CZ  CE2
BOND CB  HB1  CB HB2  CD1 HD1  CD2 HD2  CE1 HE1
DOUBLE   CD1 CG  CZ CE1   CE2 CD2
BOND CE2 HE2  CZ HZ


END

read param card flex append
* Parameters for the aminoacyl group of tRNA
```

```
*

BONDS
    CN7B     OS  241.10    1.4482

ANGLES
    CN7B    CN7B     OS   48.76  105.90  !
    CN7     CN7B     OS   48.76  105.90  !
    CN8     CN7B     OS   42.38  109.00  !
    CN7B     OS      CD   12.48  119.29  !
    HN7     CN7B     OS   10.71  113.98  !
    NH2     CT1      CD   43.68   96.89  !

DIHEDRALS
     OS     CN7B    CN7B    ON6    0.00    3       0.0
     OS     CN7B    CN7B    HN7    0.00    3     180.0  ! RNA
     CD      OS     CN7B    CN7B   0.60    3       0.0
     CD      OS     CN7B    CN8    0.60    3       0.0
     CD      OS     CN7B    HN7    0.00    3       0.0
     CD     CT1     NH2       H    0.47    3       0.0

IMPROPER

END
RETURN
```