Mechanistic insights into the slow peptide bond formation with D-amino acids in the ribosomal active site

Sergey V. Melnikov1, Nelli F. Khabibullina2, Elisabeth Mairhofer3, Oscar Vargas-Rodriguez1, Noah M. Reynolds1, Ronald Micura3,*, Dieter Söll1,4,* and Yury S. Polikanov2,5,*

1Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA, 2Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL 60607, USA, 3Institute of Organic Chemistry at Leopold Franzens University, A-6020 Innsbruck, Austria, 4Department of Chemistry, Yale University, New Haven, CT 06520, USA and 5Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, Chicago, IL 60607, USA

Received October 03, 2018; Revised November 14, 2018; Editorial Decision November 19, 2018; Accepted November 20, 2018

ABSTRACT

During protein synthesis, ribosomes discriminate chirality of amino acids and prevent incorporation of D-amino acids into nascent proteins by slowing down the rate of peptide bond formation. Despite this phenomenon being known for nearly forty years, no structures have ever been reported that would explain the poor reactivity of D-amino acids. Here we report a 3.7 Å-resolution crystal structure of a bacterial ribosome in complex with a D-aminoacyl-tRNA analog bound to the A site. Although at this resolution we could not observe individual chemical groups, we could unambiguously define the positions of the D-amino acid side chain and the amino group based on chemical restraints. The structure reveals that similarly to L-amino acids, the D-amino acid binds the ribosome by inserting its side chain into the ribosomal A-site cleft. This binding mode does not allow optimal nucleophilic attack of the peptidyl-tRNA by the reactive α-amino group of a D-amino acid. Also, our structure suggests that the D-amino acid cannot participate in hydrogen-bonding with the P-site tRNA that is required for the efficient proton transfer during peptide bond formation. Overall, our work provides the first mechanistic insight into the ancient mechanism that helps living cells ensure the stereochemistry of protein synthesis.

INTRODUCTION

All living cells, from bacteria to human, contain both L- and D-amino acids. However, only L-amino acids are utilized for protein synthesis. The mechanism of this exclusive use of L-amino acids is not yet fully understood. This is especially notable in bacteria whose cytosol contains about a dozen different D-amino acids that are used as a carbon source, signaling molecules, or building blocks for peptido-glycan cell wall synthesis (1,2). In some bacteria, D-amino acids are present in millimolar concentrations, sometimes with the levels of D-isomers exceeding those of their L-isomers (as in the case of D-alanine and D-glutamate) (2,3). In eukaryotes, nano- to micromolar concentrations of D-amino acids are typically present in animals, plants, and fungi (4,5). Thus, organisms, from bacteria to higher eukaryotes, utilize only the L-amino acids for protein synthesis despite the presence of D-amino acids in cell cytosol.

The exclusion of D-amino acids from the ribosome-dependent protein synthesis is achieved through the cooperation of at least four independent mechanisms. First, the aminoacyl-tRNA-synthetases, which select amino acids for protein synthesis, react markedly slower with D-amino acids than with L-amino acids. For instance, tyrosyl-tRNA synthetase utilizes D-tyrosine at ∼25-fold slower rate than L-tyrosine to produce tyrosyl-tRNA (6). Second, if D-aminoacyl-tRNA is formed, it is typically deacylated by the D-aminoacyl-tRNA deacylase (DTD) (7,8). This enzyme is conserved across the three domains of life and prevents accumulation and toxicities of D-aminoacyl-tRNAs (9). Third, if a D-aminoacyl-tRNA escapes hydrolysis by DTD, it is recognized by the elongation factor EF-Tu. However, its delivery to the ribosome occurs with ∼250-
fold lower yield compared to L-aminoacyl-tRNAs (10). Finally, in vitro studies showed that if a D-aminoacyl-tRNA binds the ribosomal A site, it reacts with a P-site substrate at about three orders of magnitude slower rate compared to L-aminoacyl-tRNAs, illustrating that D-amino acids markedly reduce the rate of the peptide bond formation (11). If a D-amino acid is incorporated into a nascent peptide and translocated to the P site, it might cause translation arrest, suggesting that D-amino acids also interfere with the passage of the nascent peptide through the ribosomal exit tunnel (11). Thus, cells have intricate fidelity control systems that favor preferential usage of the L-isomers over the D-amino acids at every stage of protein synthesis.

In the past years, the interest to the D-amino acid recognition by the ribosome has been revitalized due to progress in the genetic code expansion of living cells (12–14). Over the past two decades, methods have been developed that allow to genetically encode >200 non-proteinogenic amino acids to enable their ribosome-dependent and site-specific incorporation into proteins in vivo. These amino acids include: (i) post-translationally modified residues to explore the role of modifications in protein activity; (ii) photo-crosslinking side-chains to enable detection of transient protein interactions; (iii) fluorescent groups and self-labeling tags for improved live imaging; (iv) heavy atom-containing amino acids for X-ray crystallography; (v) residues with β-amino-acid backbone to endow proteins with resistance to proteolysis and others (12–15). However, all of the amino acids that have been successfully incorporated in vitro comprise only the L-isomers, while genetic encoding of D-amino acids remains a challenge.

Ribosomal synthesis of proteins with D-amino acids is desired, because site-specific replacement of L-amino acid residues with their D-isomers renders corresponding peptides protease-resistant, as it was shown for hormones and other pharmacologically active polypeptides (16–23). Also, D-amino acids are present in natural proteins (introduced via post-translational isomerization), such as bacterial lantibiotics, opioid peptides from frogs, and conotoxins (24). Therefore, the ability to perform ribosomal synthesis of D-amino acid-containing proteins is required to enable the large-scale and cost-effective production of pharmacologically active proteins and peptides.

Over the past years, messenger RNA-dependent synthesis of D-amino acid-containing proteins became possible in vitro via engineering of different translation machinery components. For example, protein engineering allowed the creation of aminoacyl-tRNA synthetases that selectively use D-isomer of tyrosine (25,26). Also, development of engineered catalytic RNAs, flexizymes, made possible production of D-aminoacyl-tRNAs for use in cell-free protein translation systems (27). Optimization of in vitro translation systems allowed synthesis of detectable amounts of peptides containing up to 10 consecutive D-amino acids (28–31). Further improvements were accomplished by random mutagenesis of ribosomal RNA (rRNA) (32,33). For instance, ribosomes carrying mutations in 2447GAUA2450 nucleotides in the 23S rRNA showed markedly improved compatibility with D-amino acids. However, these ribosome mutants were less accurate and highly toxic in Escherichia coli preventing their use in vivo (33).

The rational engineering of ribosomes to enable efficient usage of D-amino acids is currently limited due to the lack of a structural basis for the poor reactivity of D-amino acids in the peptide bond formation reaction. To overcome this limitation, we determined the crystal structure of the 70S ribosome in complex with the D-aminoacyl-tRNA mimic, ACCA-D-phenylalanine (ACCA-D-Phe), bound to the ribosomal A site. Our structure reveals that the D-aminoacyl-tRNA analog binds the ribosome in a similar fashion as L-aminoacyl-tRNAs with the CCA-end binding the ribosomal A site in a canonical way and with the D-amino acid side chain accommodated by the ribosomal side chain-binding pocket. However, due to a ‘mirror’ arrangement of the substituents at the Cα-atom of the D-amino acid, the reactive α-amino group of the D-aminoacyl-tRNA analog should deviate significantly from the optimal position that is required for the nucleophilic attack onto the carbonyl carbon of the P-site substrate. Thus, our study reconciles the observed poor reactivity of D-amino acids in ribosomal protein synthesis. The reported structure provides an essential framework for the future rational design of the PTC and its surroundings to improve the usage of D-amino acids by the ribosome.

MATERIALS AND METHODS
Synthesis of hydrolysis-resistant D-phenylalanyl-tRNA analog

The L-aminoacyl-tRNA mimic, cytidine-cytidine-puromycin (CC-Pmn) was obtained from Thermo Scientific. The D-aminoacyl-tRNA mimic, adenosine-cytidine-cytidine-adenosyl-D-phenylalanine (ACCA-D-Phe) was chemically synthesized. Each of these two tRNA analogs comprised a 3′-amido linkage between the 3′-terminal adenosine of the tRNA mimic and the C-terminus of the D-Phe or L-methyl-Tyr moieties to enable hydrolysis of the analog during crystallization. The ACCA-D-Phe conjugate was produced as outlined in Figure 1A and as described below (similar to the synthesis schemes previously reported in references (34–36)). D-phenylalanine (>98% purity) was purchased from Iris Biotech GmbH and Fluka.

1. N-(9-Fluorenyl)methoxycarbonyl-D-phenylalanine (compound 2), D-Phenylalanine 1 (0.50 g, 3.03 mmol) and Na₂CO₃ (1.76 g, 16.60 mmol) were suspended in 20 ml of 1,4-dioxane/H₂O (1/1) and cooled to 0°C. At this point, 9-fluorenylmethoxycarbonyl chloride (0.86 g, 3.32 mmol) was added to the suspension and stirred for 5 min at 0°C. The ice bath was removed and the reaction mixture was stirred for 7 h at room temperature and afterwards the reaction mixture was acidified with concentrated HCl to pH 2. The resulting solution was extracted with dichloromethane (100 ml), the organic phase was washed twice with H₂O (50 ml) and dried over Na₂SO₄. The solvent was evaporated and the crude product was purified by column chromatography on SiO₂ (CH₂Cl₂/MeOH, 100/0–96/4 v/v). Yield: 1.01 g of N-(9-fluorenyl)methoxycarbonyl-D-phenylalanine 2 as white foam (86%). TLC: (CH₂Cl₂/MeOH, 9/1): Rf = 0.60. ESI-MS (m/z) [M+H]+ calculated 388.1543; found
Figure 1. Chemical synthesis of a short hydrolysis-resistant D-phenylalanyl-tRNA analog. (A) Synthetic route. Letters indicate specific reaction conditions as follows: (a) 1.1 equivalent of Fmoc-Cl, Na₂CO₃, in 1,4-dioxane/H₂O, room temperature, 7 h, yield 86% for compound 2; (b) 1.3 equivalent of Fmoc-D-Phe, 1.3 equivalent of O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate, 1.3 equivalent of 1-hydroxybenzotriazole hydrate, 1.5 equivalent of N,N-diisopropylhexylamine in DMF, room temperature, 14 h, yield 70% for compound 3; (c) i. 2.5 equivalent of adipic acid bis(pentafluorophenyl)ester, 1 equivalent of 4-(N,N-dimethylamino)pyridine in N,N-dimethylformamide/pyridine (1/1, v/v), room temperature, 1 h, yield 52% of crude ester; ii. ~1 equivalent of amino-functionalized polystyrene support (GE Healthcare, Custom Primer SupportTM 200 Amino), pyridine, N,N-dimethylformamide, room temperature, one day, loading: 37 μmol/g for solid-support 5. (d) RNA solid-phase synthesis, deprotection, and purification. DMT – 4,4′-dimethoxytrityl, Fmoc – N(9-fluorenylmethoxycarbonyl. (B) Anion-exchange HPLC profiles of crude (top) and purified (bottom) ACCA-D-Phe conjugate. Anion-exchange chromatography conditions: Dionex DNApac PA-100 (4 × 250 mm) column; temperature: 60°C; flow rate: 1 ml/min; eluant A: 25 mM Tris–HCl (pH 8.0), 6 M urea; eluant B: 25 mM Tris–HCl (pH 8.0), 6 M urea, 500 mM NaClO₄; gradient: 0–40% B in A within 25 min; UV detection at 260 nm. (C) LC-ESI mass spectra of the purified product ACCA-D-Phe (compound 6).
19.78, 20.21 (N(CH₂CH₂CH₃)); 29.72, 30.98 (N(CH₂CH₂CH₃)); 38.71 ((C(β, Phe)); 45.24 N(CH₃CH₂CH₂CH₃)); 47.12 (C(9, Fmoc)); 51.95 N(CH₂CH₂CH₂CH₃); 52.44 (C(3′)); 56.35 (2 × O-CH₃(DMT)); 56.57 (C(α,Phe)) 63.34 (C(5′)); 67.10 (C(methylene, Fmoc)); 74.69 ((C(2′)); 83.83 (C(4′)); 86.56; 91.21 (C(1)); 113.26; 120.12; 125.14; 126.49; 126.96; 127.21; 127.34; 127.88; 129.75; 128.32; 128.96; 129.39; 130.20; 135.73; 135.75; 136.34; 139.57 (C(8′)); 141.43; 143.75; 143.89; 144.50; 150.67; 152.41; 155.97; 158.50; 158.60; 158.89; 160.53; 171.36 (C=O (Phe)) ppm. ESI-MS (m/z): [M+H]+ calculated for C₆₄H₆₉N₈O₈, 1077.5233; found 1077.5247.

3. DMTO-rA₃-NH-D-Phe-NHFmoc solid support (compound 5). To a solution of compound 4 (80 mg, 0.074 mmol) in DMF and pyridine (1.0 ml each) was added DMAP (9 mg, 0.074 mmol) and bis(pentafluorophenyl) adipate (88.7 mg, 0.19 mmol). The mixture was stirred for one hour followed by evaporation of the solvents. The residue was added to filtration over SiO₂ yielding 53 mg of the crude ester as a white foam (TLC (20% acetone in CH₂Cl₂)). The crude ester was dissolved in dry DMF (2.0 ml), and pyridine (80 µl) was added. To this solution, amino-functionalized support (GE Healthcare, Custom Primer Support™ 200 Amino, 160 mg) was added, and the suspension was agitated for 18 h at room temperature. Subsequently, the beads were collected on a Büchner funnel and washed with DMF, methanol, and CH₂Cl₂. For capping of unreacted amino groups, the beads were treated with a mixture of solution A: 25 mm Tris–HCl (pH 8.0), 6 M urea, and solution B: 25 mM Tris–HCl (pH 8.0), 0.5 M NaClO₄, 6 M urea; gradient: 0–60% B in A within 45 min or 0–40% B in A within 30 min, UV detection at λ = 260 nm.

6. Purification of the 3′-D-phenylalanyl-ACCA conjugate. The crude conjugate was purified on a semipreparative Dionex DNA Pac PA-100 column (9 × 250 mm) at 60°C with a flow rate of 2 ml/min for eluents, see above). Fractions containing the conjugate were loaded on a C18 SepPak Plus cartridge (Waters/Millipore), washed with 0.1–0.15 M (Et₃NH)₂HCO₃, H₂O, and eluted with H₂O/CH₃CN (1:1). Conjugate-containing fractions were evaporated to dryness and dissolved in H₂O (1 ml). The quality of the purified conjugate was analyzed by analytical anion-exchange chromatography (Figure 1B). The molecular weight of the synthesized conjugate was confirmed by LC-ESI mass spectrometry (Figure 1C). Yields were determined by UV photometric analysis of conjugate solutions. The final compound was dissolved in water to achieve ~50 mM concentration for stock solutions and later used in cryo-crystallization experiments.

Crystallographic structure determination
Ribosome–mRNA–tRNA complex was pre-formed by programming 5 µM 70S T7h ribosomes with 10 µM mRNA and incubation at 55°C for 10 min, followed by addition of 20 µM P-site (tRNA₆₅) and 100 µM A-site (CC-Pmn or ACCA-D-Phe) substrates (with minor changes from [37]). Each of these two steps was allowed to reach equilibrium for 10 min at 37°C in the buffer containing 5 mM HEPES–KOH (pH 7.6), 50 mM KCl, 10 mM NH₄Cl, and 10 mM Mg(CH₃COO)₂. Crystals were grown by vapor diffusion in sitting drop crystallization trays at 19°C. Initial crystalline needles were obtained by screening around previously published ribosome crystallization conditions [38–40]. The best-diffracting crystals were obtained by mixing 2–3 µl of the ribosome complexes with 3–4 µl of a reservoir solution containing 100 mM Tris–HCl (pH 7.6), 2.9% (w/v) PEG-200K, 7–12% (v/v) MPD, 100–200 mM arginine, 0.5 mM β-mercaptoethanol [41]. Crystals appeared within 3–4 days and grew up to 150 × 150 × 1600 µm in size within 10–12 days. Crystals were cryo-protected stepwise using a series of buffers with increasing MPD concentrations until reaching the final concentration of 40% (v/v) MPD, in which they were incubated overnight at 19°C. In addition to MPD,
Table 1. X-ray data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>70S complex with P-site tRNA and CC-Pmn</th>
<th>70S complex with P-site tRNA and ACCA-D-Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diffraction data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>$P2_1\overline{2}1\overline{2}1$</td>
<td>$P2_1\overline{2}1\overline{2}1$</td>
</tr>
<tr>
<td>Unit cell dimensions, Å ($a \times b \times c$)</td>
<td>212.24 × 452.84 × 620.30</td>
<td>211.39 × 452.15 × 617.64</td>
</tr>
<tr>
<td>Wavelength, Å</td>
<td>0.9795</td>
<td>0.9795</td>
</tr>
<tr>
<td>Resolution range (outer shell), Å</td>
<td>213–3.70 (3.80–3.70)</td>
<td>255–3.70 (3.80–3.70)</td>
</tr>
<tr>
<td>$I/\langle I\rangle$ (outer shell with $I/\langle I\rangle = 1$)</td>
<td>6.87 (0.88)</td>
<td>5.59 (0.86)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness (outer shell), %</td>
<td>99.4 (99.0)</td>
<td>98.5 (98.5)</td>
</tr>
<tr>
<td>$R_{merge}$ (outer shell)%</td>
<td>23.7/27.8</td>
<td>23.6/27.9</td>
</tr>
<tr>
<td>No. of crystals used</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No. of Reflections Used:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>5428 817</td>
<td>2 154 810</td>
</tr>
<tr>
<td>Unique</td>
<td>627 226</td>
<td>615 531</td>
</tr>
<tr>
<td>Redundancy (outer shell)</td>
<td>8.65 (7.95)</td>
<td>3.50 (3.14)</td>
</tr>
<tr>
<td>Wilson $B$-factor, Å$^2$</td>
<td>135.2</td>
<td>123.6</td>
</tr>
</tbody>
</table>

$R_{merge} = \Sigma I - \langle I\rangle / \Sigma I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity from multiple measurements.

$R_{work} = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$. For calculation of $R_{merge}$, 5% of the truncated dataset was excluded from the refinement.

all stabilization buffers contained 100 mM Tris–HCl (pH 7.6), 2.9% (w/v) PEG-20K, 50 mM KCl, 10 mM NH$_4$Cl, 10 mM Mg(CH$_3$COO)$_2$ and 6 mM β-mercaptoethanol. CC-Pmn or ACCA-D-Phe were not added to any of the cryo-protection solutions. After stabilization, crystals were harvested and flash frozen in a nitrogen cryo-stream at 80°C (Oxford Cryosystems).

Diffraction data were collected at the beamlines 24ID-C and 24ID-E at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). A complete dataset for each ribosome complex was collected using 0.979 Å wavelength at 100K from multiple regions of the same crystal using 0.3° oscillations. The raw data were integrated and scaled using the XDS software package (42). All crystals belonged to the primitive orthorhombic space group $P2_1\overline{2}1\overline{2}1$ with approximate unit cell dimensions of 210 Å × 450 Å × 620 Å and contained two copies of the 70S ribosome per asymmetric unit. Each structure was solved by molecular replacement using PHASER from the CCP4 program suite (43). The search model was generated from the previously published structure of the *T. thermophilus* 70S ribosome with all modifications and with bound mRNA and P-site tRNA (PDB entry 4Y4P from (41)). The initial molecular replacement solutions were refined by rigid body refinement with the ribosome split into multiple domains, followed by 10 cycles of positional and individual B-factor refinement using PHENIX (44). Non-crystallographic symmetry restraints were applied to 4 domains of the 30S ribosomal subunit (head, body, spur, helix 44), and four domains of the 50S subunit (body, L1-stalk, L10-stalk, C-terminus of the L9 protein).

Atomic models of CC-Pmn and ACCA-D-Phe were generated from their known chemical structures using PRODRG online software (45), which was also used to generate restraints for energy minimization and refinement based on idealized 3D geometry. Atomic models and restraints were used to fit/refine each of the tRNA mimics into the obtained unbiased electron density maps (Figure 2). The final model of the 70S ribosome in complex with CC-Pmn or ACCA-D-Phe and mRNA/tRNAs was generated by multiple rounds of model building in COOT (46), followed by refinement in PHENIX (44). The statistics of data collection and refinement are compiled in Table 1. All figures showing atomic models were generated using PyMol software (www.pymol.org).
RESULTS

Crystal structures of the 70S ribosome in complex with L- and D-aminoacyl-tRNA analogs in the ribosomal A site

To provide structural insights into the poor reactivity of the D-aminoacyl-tRNAs in the peptide bond formation we determined the crystal structure of *T. thermophilus* 70S ribosomes in complex with hydrolysis-resistant analogs of aminoacyl-tRNAs. We used cytidyl-cytidyl-puromycin (CC-Pmn) as an L-aminoacyl-tRNA analog that carried L-methyl-tyrosine residue (Figure 2A, L-mTyr) and adenylyl-cytidyl-cytidyl-D-phenylalanine (ACCA-D-Phe) as a D-aminoacyl-tRNA analog that carried D-phenylalanine residue (Figure 2B, D-Phe). These short analogs mimic the 3-terminal CCA-ends of the acceptor stem of full-length tRNAs. Both analogs carried a D- or L-amino acid attached to the 3'-terminal nucleotide via the amide linkage (instead of the naturally occurring ester bond) to prevent spontaneous hydrolytic deacylation during the crystallization (Figure 2A, B). Using either of these compounds as an A-site substrate, we determined their crystal structures in complex with the *T. thermophilus* 70S ribosomes carrying messenger RNA and tRNA<sub>Met</sub> in the P site (Materials and Methods). Although the P-site tRNA in both of our complexes was represented by the deacylated tRNA<sub>Met</sub>, which is not strictly physiological, previous studies have shown that aminoacylation status of the P-site tRNA does not affect conformation of the amino acid attached to the A-site tRNA substrate (37,47-49). Therefore, it is reasonable to assume that the conformation and interactions of the A-site substrates in our structures are identical to those seen in physiologically more relevant complexes of the ribosome.

Both crystal structures were determined at 3.7 Å resolution by molecular replacement using the atomic coordinates of the *T. thermophilus* 70S ribosome with the A-site tRNA removed (PDB entry 4Y4P (41)) (Table 1). The unbiased difference Fourier maps revealed unique positive den-
Figure 3. D-aminoacyl-tRNA analog establishes canonical A-loop interactions. Watson-Crick base-pairing between the penultimate cytidine of the (A) L-Phe-tRNAPhe (green, PDB entry 1VY4 (37)) or (B) D-aminoacyl-tRNA analog ACCA-D-Phe (yellow) and the nucleotide G2553 in the A-loop (Helix 92) of the 23S rRNA (light blue). Note that these A-loop interactions play a key functional role in accommodation and proper positioning of the substrate in the A site of the ribosome.

CCA-end of the D-aminoacyl-tRNA manifests canonical interactions with the ribosome

Next, we assessed whether the observed poor reactivity of D-aminoacyl-tRNAs might be caused by suboptimal positioning of the CCA-end in the A site. In the case of canonical L-aminoacyl-tRNAs, their binding to the A site results in specific interactions of the CCA-end with the A-loop (Helix 92) of the 23S rRNA, in which the tRNA residue C75 forms Watson-Crick base-pair with the nucleotide G2553 (here and throughout the text we use *E. coli* numbering of rRNA nucleotides) (47). This interaction is required for the proper positioning of the aminoacyl-tRNA substrate in the PTC (47). Our electron density maps revealed that, in both crystal structures, the CCA-ends of aminoacyl-tRNA analogs establish canonical contacts with the A-loop (Figure 3B), illustrating that the presence of D amino acid residue does not impede recognition of the CCA-end by the ribosome (Figure 2E, F).

Side chains of L- and D-amino acid residues bind the A site in a similar fashion

We next asked how D-amino acid binding to the A site is compared to that of L-amino acids. The electron density maps revealed the backbones of both L- and D- amino acids, as well as the entire side chain of the L-methyltyrosine (Figures 2C and 4C, D). The side chain of the D-phenylalanine (Figures 2D and 4E, F) was also visible in the electron density map, although only up to the Cγ atom, pointing to the partial flexibility of the D-amino acid side chain. The position of the Cγ atom of the D-phenylalanine side chain and the fact that the bulky side chain of D-Phe should tightly fit into the A-site cleft indicates that the D-Phe side chain is fully accommodated into the A-site cavity because any other orientation of the tip of D-Phe side...
chain would result in collisions with the surrounding rRNA nucleotides (Figure 4C–F). Thus, our data indicate that not only the CCA-end of the D-aminoacyl-tRNA analog forms canonical interactions with the ribosome but also the D-amino acid side chain binds the A-site cleft in a fashion similar to that of the L-amino acids side chains before the peptide bond formation takes place.

D-amino acid adopts a poorly reactive conformation in the ribosomal A site

Next, we checked whether binding of the D-Phe side chain to the A-site cleft affects the conformation of its reactive α-amino group. Of course, at the 3.7Å resolution, it is impossible to identify the exact location of the α-amino group, although this is also impossible even at higher resolutions at which ribosome structures were reported previously (such as 2.3–2.5Å). What is crucial here is the fact that we observe electron density for the part of the D-Phe side chain. And by applying ideal chemical restraints, from the known locations of the Cα, Cβ and Cγ atoms we can unequivocally deduce the location of the reactive α-amino group and compare it with that of L-mTyr (Figure 5A).

In case of L-amino acids, the α-amino group (such as the one in CC-Pmn) forms a hydrogen bond (H-bond) with the 2′-hydroxyl of the nucleotide A76 of the P-site tRNA (Figure 4C, black dotted line), which plays critical role in (i) its positioning for the optimal nucleophilic attack onto the carbonyl carbon of the P-site substrate and (ii) the subsequent shuttling of the protons (Figure 5A, solid arrowhead) (37). By contrast, in our structure of ribosome-bound ACCA-D-Phe substrate, the α-amino group of the D-amino acid is directed not towards, but away from the P-site tRNA (Figure 5A). In this orientation, the α-amino group is unlikely to form such H-bond due to the at least 3.5Å distance between the α-amino group and the 2′-OH group of the P-site tRNA and unfavorable geometry (Figures 4E and 5A). As a result, the nucleophilic attack from this position is likely to be less efficient due to the larger distance between the reactive α-amino group and the carbonyl carbon of the P-site substrate (Figure 5A, dashed arrowhead). Also, a defined path for abstraction of the proton from the attacking α-amino is not arranged (37), which might be another reason for the de-
Figure 5. Reactive conformation of the D-amino acid is likely prevented by the conserved rRNA residues in the peptidyl-transferase center. (A) Comparison of the observed structures of the L- and D-aminoacyl-tRNA analogs bound to the ribosome. Shown are the 23S-rRNA-aligned energy-minimized conformations of CC-Pmn (magenta) and ACCA-D-Phe (yellow) bound to the ribosomal A site. Note that, due to the opposite chirality of the C/H9251-atoms, the C/H9251-amino group of the D-phenylalanine is positioned further away from the carbonyl carbon of the P-site substrate resulting in reduced reactivity. (B) The conformation of the D-phenylalanine, in which its α-amino group is aligned for the optimal nucleophilic attack onto the carbonyl carbon of the P-site substrate. In this state, the side chain of D-amino acid (especially C/H9252-atom) severely clashes with the key functional nucleotide U2506 of the PTC. (C) Mutations in the 23S rRNA that improve ribosomal usage of D-amino acids. Relative locations of the 23S rRNA residues A2451 and C2452 forming the A-site cleft (light blue spheres) and the residues G2447, A2448, U2449 and A2450, whose mutations improve utilization of the D-amino acids by the ribosome (blue). Shown is the close-up view of the PTC with bound P-site tRNA (dark blue) and A-site short substrate ACCA-D-Phe (yellow). Note that residues G2447 and A2450 (blue spheres) are located near the A-site cleft. Mutations of these purine residues to smaller pyrimidines might lead to either an increased size of the A-site cleft or increased flexibility of the adjacent residues forming the A-site cleft.

The role of the A-site cleft in the amino acid positioning in the A site

A-site cleft binds the side chains of incoming amino acids and plays a critical role in the positioning of the incoming amino acids in the PTC. Previously, all the observed L-amino acid residues were shown to adopt a uniform position in the ribosomal A site, where the Cβ-atom (and all other atoms) of the side chain is always directed into the A-site cleft (37,47,48,50,51). Such orientation was suggested to help physically exclude amino acid side chain from the catalytic center of the ribosome and, thereby, prevent potential steric clashes and non-desired chemical reactivities of amino acid side chains in the peptide bond formation. Our observation that the D-amino acid side chain accommodates into the A-site cleft suggests that this conserved component of the ribosomal catalytic center might also be involved in the stereospecificity control of protein synthesis.

Key determinants of L- versus D-amino acid discrimination

Previous studies suggested two alternative models of how ribosome can discriminate between the two possible chiralities of the incoming amino acids and reject D-amino acids from the use in protein synthesis. One model, based on the molecular modeling attempts using pioneering structures of archaeon Halocarcha marismortui, assigned the critical role in rejecting D-amino acids to the nucleotide U2585 in the 23S rRNA (52,53). In another model, based on structural analysis of the pre-attack state of the peptide bond formation reaction, the critical role in discriminating amino acid chirality has been assigned to the nucleotide U2506 (54). The main difference between these two models stemmed from the lack of knowledge about the orientation of the D-
Insights into the mechanism of poor D-amino acid reactivity in the ribosomal A site

Our structure suggests that the poor reactivity of the D-aminoacyl-tRNAs stems from the suboptimal positioning of the reactive \(\alpha\)-amino group, which hinders the nucleophilic attack and blocks the proton shuttle during the peptide bond formation (Figure 5A). As suggested earlier, this suboptimal orientation of the \(\alpha\)-amino group of the D-amino acid substrate appears to be caused by the residue U2506 in the PTC, which does not allow such substrate to adopt optimal conformation for the nucleophilic attack (54).

Also, our study might explain previous findings that the use of D-amino acids by the ribosome can be improved through the mutagenesis of the 23S rRNA (32,33). For example, mutations of the \(2447\)GAUA\(2450\) segment of the \(E.\) \(coli\) 23S rRNA to \(2447\)UGGU\(2450\) or \(2447\)UGGC\(2450\) lead to 5-fold improvement of D-phenylalanine or D-methionine incorporation into a reporter protein \(in\) \(vitro\) (32,33). The mutated \(2447\)GAUA\(2450\) segment of the 23S rRNA is located in the ribosomal A site, with residues G2447 and A2450 being in the direct vicinity of the A-site cleft (Figure 5C). Mutations of the purine nucleotides G2447 and A2450 to smaller pyrimidines should increase the size of the A-site cleft, thereby allowing the D-amino acids to adopt more reactive conformations without clashing of their side chains with the residues of the PTC.

Implications for the ribosome engineering

Over the past two decades, ribosome engineering produced an array of ribosome variants for applications in basic research and biotechnology (55). Ribosomes with mutated anti-Shine-Dalgarno sequence (56) and with tethered ribosomal subunits (57,58) were constructed to allow the presence of two independent translation systems in a single cell. Hybrids between bacterial and eukaryotic ribosomes were constructed to explore principles of antibiotic specificity (59) or use bacterial ribosomes in eukaryote-derived \(in\) \(vitro\) translation systems (60). Also, engineered ribosomes were produced to decode quadruplet codons (61) or recognize artificial tRNAs (62). Finally, ribosome variants were constructed to improve ribosome compatibility with non-canonical amino acids, such as D-amino acids (32,33) and \(\beta\)-amino acids (63,64).

By showing D-amino acid residue in the ribosomal catalytic center, our structure provides the basis for the rational design of the amino acid binding pocket to improve ribosome compatibility with non-canonical substrates. In this regard, it is important to note that ribosomes with altered A-site cleft have been previously observed in nature. First, the structure of the A-site cleft was shown to vary across species. In eukaryotes, the 80S ribosomes carry C2452U substitution that makes the A-site cleft slightly larger in eukaryotes compared to bacteria and archaea (51). Also, mutations in the A-site cleft or its vicinity are present in mitochondrial ribosomes. In yeast, mitochondrial ribosomes carry G2447A substitution that alters the A-site cleft structure and confers resistance to the A-site targeting antibiotic chloramphenicol, and in mice, mitochondrial ribosomes carry the A2451U substitution that also confers chloramphenicol resistance (65–67). Apart from natural variations of the A site structure, ribosomes with altered A-site cleft were produced in the laboratory by mutagenesis of ribosomal protein L3 and the 23S rRNA residue G2447 (68). Importantly, these mutations are not lethal and only moderately affect the efficiency of protein synthesis (68). Moreover, the impact of the A-site cleft structure on the peptide bond formation was predicted by previous studies employing either theoretical quantum mechanics (69) or molecular dynamics simulations (70). Collectively, previous studies and our structural analyses suggest that the described above ribosome variants with altered A-site cleft might be good candidates for the development of \(in\) \(vitro\) compatible and chirality-promiscuous engineered ribosomes.

In addition to the A-site cleft mutations, a number of alterations have been explored in the ribosomal catalytic center, PTC. Although the majority of the mutations in the PTC are lethal \(in\) \(vitro\), many of them were studied \(in\) \(vitro\) where they were shown to have either moderate or no effect on protein synthesis. For instance, the use of chemically modified tRNAs for \(in\) \(vitro\) ribosome assembly allowed to introduce highly toxic and lethal mutations or artificial nucleotides into the rRNA (71,72). Using this approach, it was demonstrated that apurinization of the residue U2506 not only preserves catalytic properties of the ribosome but results in a two-fold increase of the yield of \(in\) \(vitro\) protein synthesis (72). Given that this residue prevents the reactive conformation of a D-aminoacyl-tRNA in the A site, its apurinization might improve utilization of the D-amino acids by the ribosome.

The ultimate understanding of the bases for slow incorporation of D-amino acids during ribosomal protein synthesis requires additional structural studies that will illuminate intermediate steps of peptide bond formation and nascent peptide folding in the ribosomal tunnel. Complementary biochemical and microbiological experiments could confirm the predictive power of the structures and create a new generation of synthetic ribosomes for efficient protein synthesis utilizing D-amino acids. While additional research is needed to address these questions, our current study provides the first structural insight into the ancient mechanism by which the ribosome ensures the stereospecific synthesis of natural proteins.

DATA AVAILABILITY

Coordinates and structure factors were deposited in the RCSB Protein Data Bank with accession code 6N9E for the \(T.\) \(thermophilus\) 70S ribosome in complex with CC-puromycin, mRNA, and P-site tRNA, and 6N9F for the \(T.\) \(thermophilus\) 70S ribosome in complex with ACCA-D-Phe, mRNA, and P-site tRNA.
ACKNOWLEDGEMENTS

We thank all members of the R.M., D.S. and Y.S.P. laboratories for discussions and critical feedback. We thank Hui Si Kwok and Kyle Hoffman for critical reading of the manuscript and valuable suggestions. We thank the staff at NE-CAT beamline 24ID-C for help with data collection and freezing of the crystals, especially Drs. Kanagalakhatta Rajashankar, Malcolm Capel, Frank Murphy, Igor Kourinov, Anthony Lynch, Surajit Banerjee, David Neau, Jonathan Schuermann, Narayanasami Sukumar, James Withrow, Kay Perry and Cyndi Salbego.

FUNDING

Illinois State startup funds [to Y.S.P.]; U.S. National Institutes of Health [GM122560 to D.S.]; Austrian Science Fund FWF [P27947 to R.M.]. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health [P41-GM103403 to NE-CAT]. The Pilatus 16M detector on 24-ID-E beamline is funded by a NIH-ORIP HEI [S10-RR029205 to NE-CAT]. The Eiger 16M detector on 24-ID-E beamline is funded by a NIH-ORIP HEI grant [S10-OD021527 to NE-CAT]. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357, Funding for open access charge: Illinois State startup funds. Conflict of interest statement. None declared.

REFERENCES
