

The mechanism of discriminative aminoacylation by isoleucyl-tRNA synthetase based on wobble nucleotide recognition

Received: 12 June 2024

Accepted: 27 November 2024

Published online: 30 December 2024

 Check for updates

Bingyi Chen^{1,2}, Fang Yi^{1,2}, Zhiteng Luo^{1,2}, Feihu Lu^{1,2}, Hongwei Liu³, Siting Luo^{1,2}, Qiong Gu¹ & Huihao Zhou^{1,2}✉

The faithful charging of amino acids to cognate tRNAs by aminoacyl-tRNA synthetases (AARSS) determines the fidelity of protein translation. Isoleucyl-tRNA synthetase (IleRS) distinguishes tRNA^{Ile} from tRNA^{Met} solely based on the nucleotide at wobble position (N34), and a single substitution at N34 could exchange the aminoacylation specificity between two tRNAs. Here, we report the structural and biochemical mechanism of N34 recognition-based tRNA discrimination by *Saccharomyces cerevisiae* IleRS (ScIleRS). ScIleRS utilizes a eukaryotic/archaeal-specific arginine as the H-bond donor to recognize the common carbonyl group (H-bond acceptor) of various N34s of tRNA^{Ile}, which induces mutual structural adaptations between ScIleRS and tRNA^{Ile} to achieve a preferable editing state. C34 of unmodified tRNA^{Ile}(CAU) (behaves like tRNA^{Met}) lacks a relevant H-bond acceptor, which disrupts key H-bonding interactions and structural adaptations and suspends the ScIleRS·tRNA^{Ile}(CAU) complex in an initial non-reactive state. This wobble nucleotide recognition-based structural adaptation provides mechanistic insights into selective tRNA aminoacylation by AARSS.

The genetic code is established by the faithful charging of transfer ribonucleic acids (tRNAs) with cognate amino acids, a process catalyzed by a family of ancient enzymes named aminoacyl-tRNA synthetases (AARSS)^{1,2}. In the universal genetic code table, 61 codons encode 20 common proteogenic amino acids. Generally, two pyrimidine (Y)-ending codons (NNU and NNC) encode the same amino acids, as do the two purine (R)-ending codons (NNA and NNG). However, an essential exception is that two AUR codons separately encode two different amino acids: AUA for L-isoleucine (L-Ile) and AUG for L-methionine (L-Met). In addition to AUA, L-Ile is also encoded by two AUY codons (AUU and AUC), making L-Ile the only

amino acid encoded by three codons. Like most AARSS, isoleucyl-tRNA synthetase (IleRS) recognizes anticodon triplets as the primary identity element of its tRNA substrates, and in particular, it discriminates tRNA^{Ile} from tRNA^{Met} by relying exclusively on the first (wobble) anticodon nucleotide (N34 of tRNA)^{3–5}. tRNA^{Met} can be efficiently isoleucylated by *Saccharomyces cerevisiae* IleRS (ScIleRS) when its C34 is replaced by G34⁴. Therefore, IleRS must strictly exclude C34 to avoid mis-aminoacylation of tRNA^{Met} and accommodate N34 with different sizes and chemical structures in tRNA^{Ile} isoacceptors, which is a more challenging task than that encountered by all other AARSS.

¹State Key Laboratory of Anti-Infective Drug Discovery and Development, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China. ²Guangdong Provincial Key Laboratory of Chiral Molecule and Drug Discovery, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China. ³Department of Laboratory Medicine, The Affiliated Qingyuan Hospital (Qingyuan People's Hospital), Guangzhou Medical University, Qingyuan 511518, China. ✉e-mail: zhuihao@mail.sysu.edu.cn

Various post-transcriptional modifications have been developed on the N34s of tRNA^{Ile} isoacceptors to facilitate their recognition by IleRS (Supplementary Fig. 1). In eukaryotes, modifications of tRNA^{Ile}(AAU) and tRNA^{Ile}(UAU) generate tRNA^{Ile}(IAU) and tRNA^{Ile}(ΨAUΨ), respectively (I: inosine, Ψ: pseudouridine)^{4,6,7}. The I34 modification by tRNA adenosine deaminase increased the isoleucylation of the in vitro transcript of tRNA^{Ile}(AAU) by 16-fold, while the tRNA^{Ile}(ΨAUΨ) extracted from yeast cells was 40-fold more active than the unmodified in vitro transcript of tRNA^{Ile}(UAU)⁴. Interestingly, prokaryotes encode a tRNA^{Ile} bearing the CAU anticodon, and the tRNA^{Ile}(CAU) lacking modification at C34 behaves like a tRNA^{Met}^{8,9}. The post-transcriptional modifications of C34 to 2-lysylcytidine (L34) in bacteria⁸ and to 2-agnatinylycytidine (agm²C34) in archaea⁹ efficiently changed the amino acid-accepting and mRNA-decoding specificities of tRNA^{Ile}(CAU) from L-Met to L-Ile.

To facilitate recognition of the modified or unmodified N34, IleRS recruits new C-terminal domains in addition to the canonical anticodon binding domain (ABD)⁵. C-terminal truncated IleRS is active for L-Ile activation but inactive for L-Ile tRNA transfer^{10,11}, validating the important role of the C-terminal domains in tRNA recognition. Interestingly, IleRSs from the three domains of life diverge sharply in their C-terminal domains (Fig. 1a)¹². Eukaryotic and archaeal IleRSs contain C-terminal sequences that are at least twice as long as that of bacterial

IleRS, and IleRS in higher eukaryotes has an additional unique domain (UNE-I) for multi-synthetase complex (MSC) assembly¹³. Notably, the zinc-binding domain (ZBD), which is essential for N34 recognition by bacterial IleRS⁵, does not exist in eukaryotic and archaeal IleRSs, suggesting that eukaryotic and archaeal IleRSs develop distinct N34 recognition mechanisms which may be related to eukaryotic/archaeal-specific N34 modifications.

Moreover, how a small difference at the single nucleotide N34 could control whether eukaryotic/archaeal IleRS aminoacylates a tRNA remains a mystery. The local geometry difference resulting from a single substitution at the G3-U70 base pair of tRNA^{Ala} could be transmitted along the acceptor stem, finally causing the tRNA 3' CCA end to fold back into a non-reactive route¹⁴. However, the distance between N34 and the CCA end of tRNA^{Ile} (approximately 70 Å) is significantly greater than that between G3-U70 and the CCA end of tRNA^{Ala}, and the transmission of N34 recognition information could be even more difficult.

Here, we report two parallel cocrystal structures of ScIleRS bound with tRNA^{Ile}(GAU) and unmodified tRNA^{Ile}(CAU) (Fig. 1b). The unique C-terminal domains of ScIleRS were found to employ a eukaryotic/archaeal-specific and more robust N34 recognition mechanism based on a conserved arginine. Importantly, because of the lack of the arginine-mediated N34 interactions, the ScIleRS-tRNA^{Ile}(CAU) complex

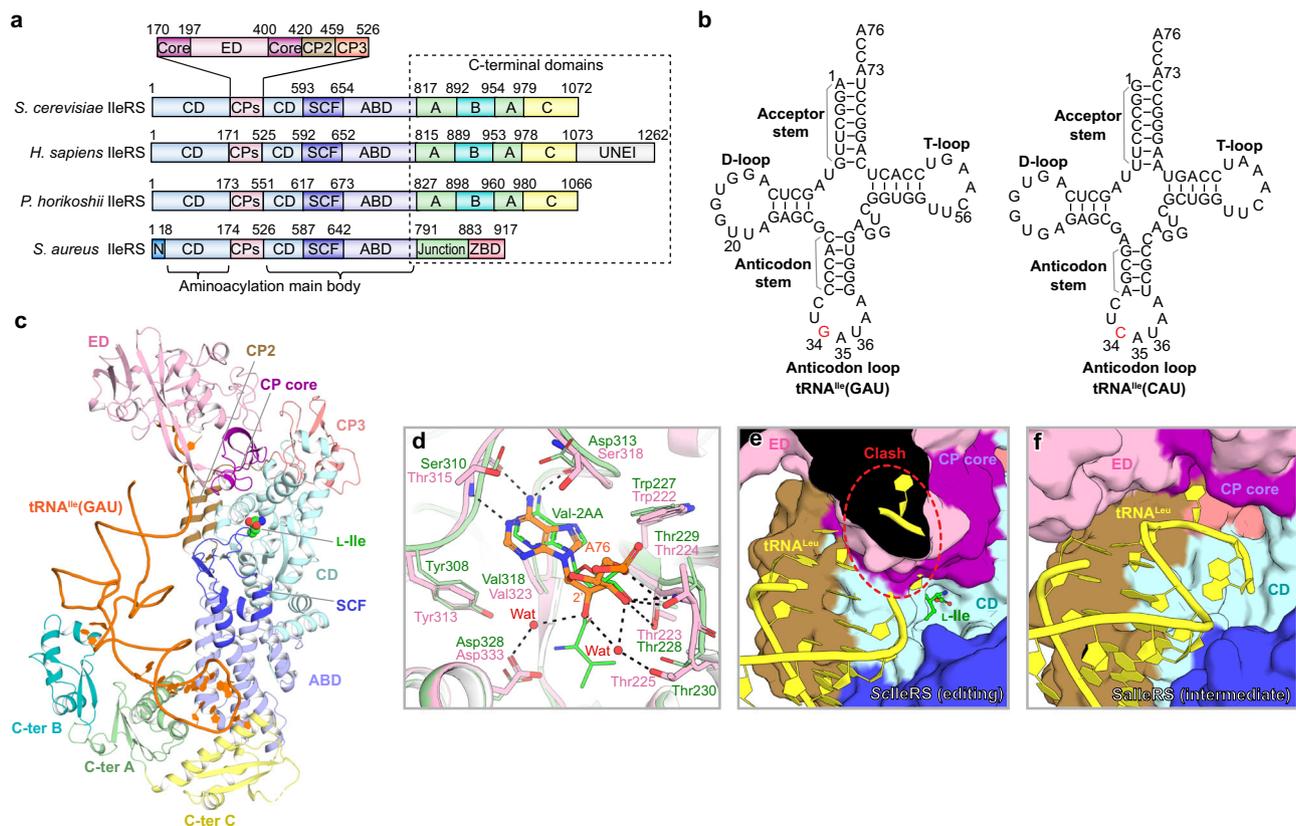


Fig. 1 | The ScIleRS-tRNA^{Ile}(GAU)-L-Ile complex structure in the editing state. a Domain diagram of IleRSs from *Saccharomyces cerevisiae*, *Homo sapiens*, *Pyrococcus horikoshii* and *Staphylococcus aureus*. **b** Cloverleaf models of tRNA^{Ile}(GAU) and tRNA^{Ile}(CAU) from *Escherichia coli*. **c** Cartoon representation of the overall structure of ScIleRS in complex with tRNA^{Ile}(GAU) and L-Ile. ScIleRS is colored the same as the domain diagram, the substrate L-Ile is represented as spheres, and tRNA^{Ile}(GAU) is colored in orange with the nucleotides directly interacting with ScIleRS shown in filled rings. **d** Structural superposition of the ScIleRS-tRNA^{Ile}(GAU)-L-Ile complex and the *Thermus thermophilus* IleRS (TtIleRS) ED-Val-2AA complex (PDB ID: 1WNZ,

green) confirms the protein-tRNA interactions in the editing state revealed by the small-molecule probe Val-2AA. The H-bonds between residues of ScIleRS and nucleotide A76 of tRNA^{Ile}(GAU) are shown as black dashed lines. **e** tRNA^{Ile} was modeled to IleRS according to the *E. coli* IleRS-tRNA^{Ile}-Leu-AMP complex structure in the aminoacylation state (PDB ID: 4AQ7). The CP core of ScIleRS in the editing state largely overlaps with the acceptor arm of tRNA^{Ile}. **f** In contrast, structural modeling revealed that there is no significant conflict between the acceptor arm of tRNA^{Ile} and the CP core of SallleRS in the SallleRS-tRNA^{Ile}(CAU)-mupirocin complex (PDB ID: 1FFY).

cannot trigger necessary structural adaptations to reach the preferable conformation that the functional ScleRS-tRNA^{Ile}(GAU) complex adopts.

Results

ScleRS and tRNA^{Ile}(GAU) form an editing state complex

Our attempts to crystallize the complex of ScleRS with *S. cerevisiae* tRNA^{Ile}(IAU) were unsuccessful. It has been reported that *Escherichia coli* tRNA^{Ile}(GAU) is recognized and isoleucylated by ScleRS with an efficiency similar to that of SctRNA^{Ile}(IAU)⁴. Consequently, *Ect*tRNA^{Ile}(GAU) was utilized as a substitute for *S. cerevisiae* tRNA^{Ile} in the study of the eukaryotic tRNA recognition mechanism. The structure of full-length ScleRS in complex with in vitro transcribed *Ect*tRNA^{Ile}(GAU) and L-isoleucine (L-Ile) (Fig. 1c) was determined by X-ray crystallography to 2.8 Å (Supplementary Fig. 2 and Supplementary Table 1). The crystallographic asymmetric unit contains two ScleRS-tRNA^{Ile}(GAU)-L-Ile ternary complexes that adopt similar conformations to each other (Supplementary Fig. 2), and the complex consisting of chains A (ScleRS) and T (tRNA^{Ile}) is discussed below owing to its superior electron density.

ScleRS comprises three parts: the aminoacylation main body, the connective peptides (CPs), and the C-terminal appendant domains contributing to tRNA binding (Fig. 1a, c). The aminoacylation main body can be further divided into the Rossmann-fold catalytic domain (CD), stem-contact fold (SCF) and anticodon-binding domain (ABD). CPs contain the CP core, editing domain (ED, also known as CP1), CP2 and CP3. Unlike bacterial IleRS, whose C-terminal domains consist of a C-terminal junction domain and a ZBD, the C-terminal sequence of ScleRS can be divided to three small domains, named the C-ter A, C-ter B and C-ter C domains (Fig. 1a, c). The C-ter A domain resembles the C-terminal junction domain of bacterial IleRS (Supplementary Fig. 3). The C-ter B domain is unique to eukaryotic and archaeal IleRSs and has no corresponding domain in bacterial IleRS. Although the C-ter C domain is the functional counterpart of the ZBD of bacterial IleRS, its structure is distinct from that of the ZBD, but despite the low sequence homology, it resembles the C-ter A domain (Supplementary Fig. 3).

The anticodon stem-loop of tRNA^{Ile}(GAU) is clamped between the aminoacylation main body and C-terminal domains (Fig. 1c). The acceptor stem of tRNA^{Ile}(GAU) forms only a few interactions with the CD, and the amino acid-accepting 3' CCA end is directed into the ED, indicating that the complex was crystallized in the editing state (Fig. 1c). Owing to their similar sizes and physicochemical properties, L-valine (L-Val) and non-proteinogenic norvaline are incorrectly activated and charged to tRNA^{Ile} by IleRS at certain rates, and IleRS quickly hydrolyzes the mischarged tRNA^{Ile} with its ED to maintain protein translation fidelity^{15,16}. The A76 of tRNA^{Ile}(GAU) interacts with the editing pocket in a manner similar to 2'-(L-valyl)amino-2'-deoxyadenosine (Val-2AA), a post-transfer editing substrate analog that mimics the 3' end of the aminoacyl-2'-ester Val-tRNA^{Ile}¹⁷ (Fig. 1d). The 2'-OH of A76 of tRNA^{Ile}(GAU) is precisely oriented towards a subpocket for the L-Val (Fig. 1d). Our ScleRS-tRNA^{Ile}(GAU)-L-Ile complex is the first structure that clearly shows how the entire 3' CCA end interacts with the ED (Supplementary Fig. 4), and it strongly supports the editing substrate recognition mechanism of IleRS suggested previously by small-molecule probes¹⁷.

Class Ia AARSs containing EDs (IleRS, LeuRS and ValRS) prefer to bind substrate tRNAs in the editing state, and thus far, the aminoacylation state complex has been captured only for LeuRS^{5,18,19}. In the editing state of the ScleRS-tRNA^{Ile}(GAU)-L-Ile ternary complex, the CP core of ScleRS packs over the L-Ile pocket, and structural comparison suggested that the CP core closes the cleft for the tRNA^{Ile} CCA end to enter the aminoacylation site (Fig. 1e). The closed conformation of the CP core has also been observed in tRNA-free *Thermus thermophilus* IleRS (TtIleRS, PDB ID: 1ILE)²⁰, *Candida albicans* IleRS (CalleRS, PDB ID: 6LDK)¹⁰ and *Helicobacter pylori* IleRS (HpIleRS, PDB ID:

8WNF)²¹ (Supplementary Fig. 5), as well as LeuRS and ValRS in tRNA-free and tRNA-editing states^{18,22,23}. Translocation of the tRNA 3' CCA end to CD requires a large rotation of the CP core and ED in these three AARSs¹⁹. In the SalleRS-tRNA^{Ile}(GAU)-mupirocin complex (PDB ID: 1FFY), although the acceptor arm of tRNA^{Ile}(GAU) is orientated towards the ED, the ED and CP core of *Staphylococcus aureus* IleRS (SalleRS) rotated by approximately 42° relative to those of ScleRS in the ScleRS-tRNA^{Ile}(GAU)-L-Ile complex (Supplementary Fig. 5), opening the conformation for translocating the tRNA 3' CCA end between the CD and ED (Fig. 1f)⁵. Thus, the structure of the SalleRS-tRNA^{Ile}(GAU)-mupirocin complex may represent an intermediate state from the editing state to the aminoacylation state, probably induced by the co-binding of the Ile-AMP-mimicking inhibitor mupirocin. In contrast, our structure provides the first unambiguous editing conformation for studying the catalytic process and possibly inhibitors of IleRS.

Productive tRNA^{Ile} binding induces C-terminal domain movements

The C-ter B domain, which is highly dynamic in tRNA-free ScleRS (PDB ID: 7D5C)¹¹, is stabilized upon tRNA^{Ile} binding through interactions with the tRNA^{Ile} elbow (consisting of D and T loops) (Fig. 1c). The C-ter B domain consists of three small α -helices, a short β -hairpin and an antiparallel two-stranded β -sheet, and the helix-turn-helix sequence from Trp900 to Ser920 is rich in basic residues (Supplementary Fig. 6). Residues Trp900, Pro901 and Lys915 form stacking, hydrophobic and hydrogen-bonding (H-bonding) interactions with nucleotides G19 and U20 of tRNA^{Ile}(GAU), respectively (Fig. 2a). As a result, U20 flips approximately 180° compared to that of tRNA^{Ile}(GAU) bound to SalleRS (PDB ID: 1FFY) (Fig. 2b). When the C-ter B domain was deleted by replacing the sequence from Val897 to Asn948 with the linker of -GSGS-, ScleRS Δ CB could still activate amino acid (Fig. 2c) but completely lost aminoacylation activity against tRNA^{Ile} (Fig. 2d), indicating that C-ter B domain deletion does not affect ScleRS folding but disrupts functional ScleRS-tRNA^{Ile} binding. Consistently, the formation of ScleRS Δ CB-tRNA^{Ile}(GAU) complex is weaker than that of the wild-type ScleRS-tRNA^{Ile}(GAU) complex as indicated by the electrophoretic mobility shift assay (EMSA) (Fig. 2e). We also introduced mutations at the G18-U55 and G19-C56 tertiary base pairs located in the elbow region of tRNA^{Ile}(GAU), and the results indicated that G19C, G18C&C56A and U55A mutants exhibited significant reductions in isoleucylation compared to the wild-type tRNA^{Ile}(GAU) (Fig. 2f). These tRNA mutants, along with ScleRS Δ CB, underscored the critical role of the C-ter B-elbow interaction in the isoleucylation of tRNA^{Ile} by ScleRS. Notably, U20, which directly interacts with the C-ter B domain, as well as the tertiary base pairs G18-U55 and G19-C56, which are important for maintaining the elbow conformation, are conserved across all tRNA^{Ile} isoacceptors in both *E. coli* and *S. cerevisiae*, suggesting that the C-ter B domain of ScleRS likely employs a similar mechanism to recognize the elbow of *S. cerevisiae* tRNA^{Ile}.

Notably, the C-ter B domain exists only in eukaryotic and archaeal IleRSs among all class I AARSs, and its best structural homolog is the insertion 3 domain (Ins3) of eukaryotic α_2 glycyI-tRNA synthetase (GlyRS, a class II AARS) (Supplementary Fig. 3), as revealed by DALI^{24,25}. The Ins3 of α_2 GlyRS as well as a similar B2 domain in bacterial $\alpha_2\beta_2$ GlyRS were proposed to undergo a large conformational movement to interact with the tRNA^{Gly} elbow, which may contribute to protecting tRNA^{Gly} from undesired disassociation during aminoacylation²⁵⁻²⁸.

When tRNA^{Ile}(GAU)-bound ScleRS was aligned with tRNA-free ScleRS (PDB ID: 7D5C) based on the CD, the ABD and C-ter A domain underwent the rotations of approximately 7° and 25°, respectively (Fig. 2g). An apo structure of full-length ScleRS was predicted by AlphaFold2 (AlphaFold DB: AF-P09436-F1)²⁹, and its ABD and C-ter A domain were well aligned with those of tRNA-free ScleRS, highlighting the reliability of the C-terminal conformation of the predicted

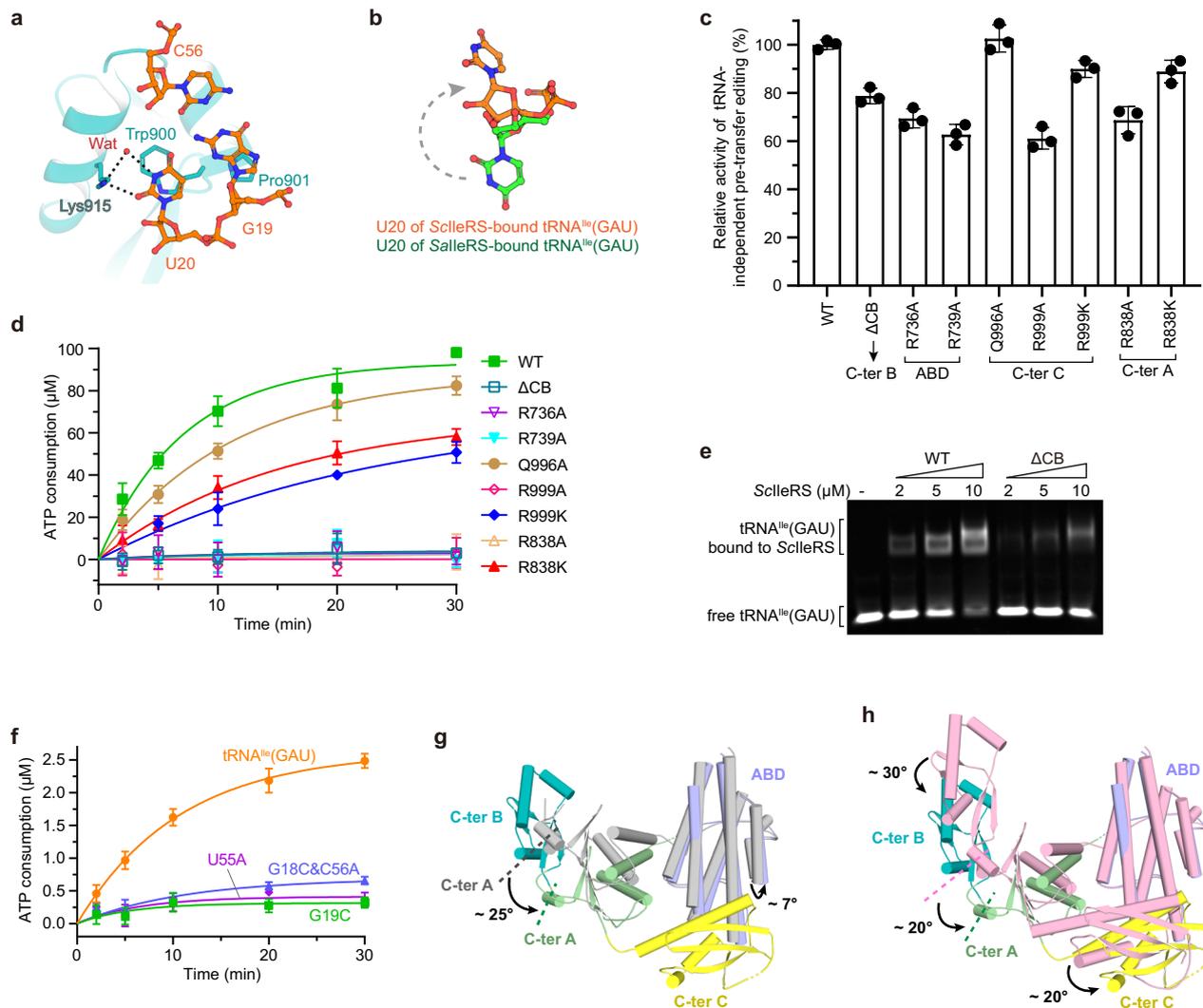


Fig. 2 | Productive tRNA^{Ile} binding requires conformational movements at ScleRS C-terminal domains. **a** The C-ter B domain of ScleRS is stabilized and binds to the elbow of tRNA^{Ile}(GAU). **b** Interactions with the C-ter B domain induce a conformation of the U20 nucleotide of ScleRS-bound tRNA^{Ile}(GAU) opposite to that of SallerRS-bound tRNA^{Ile}(GAU) (PDB ID: 1FFY). **c** ScleRS proteins with mutations located far from the active site exhibited similar or comparable activity to that of the wild-type protein in the tRNA-independent pre-transfer editing assay. Data are presented as means \pm SD ($n = 3$ independent experiments). **d** Most ScleRS variants partially or completely lost the tRNA^{Ile} isoleucylation activity as measured by tRNA-dependent ATP consumption assay. *EctRNA^{Ile}(GAU)* overexpressed in *E. coli* cells was utilized in this assay. Data are presented as means \pm SD ($n = 3$

independent experiments). **e** The EMSA assay result revealed that tRNA^{Ile} binding ability of ScleRS Δ CB is weaker than that of wild-type ScleRS. The in vitro transcript of tRNA^{Ile}(GAU) was used in this assay. Similar results were observed in two independent experiments. **f** The aminoacylation activity of ScleRS against in vitro transcribed tRNA^{Ile}(GAU) and its variants with G19C, G18C&C56A or U55A mutations. All the variants presented significantly lower isoleucylation than that of the wild-type tRNA^{Ile}(GAU). Data are presented as means \pm SD ($n = 3$ independent experiments). **g, h** Structural comparison of the tRNA^{Ile}(GAU)-bound ScleRS with the tRNA-free ScleRS (PDB ID: 7D5C, colored in gray) (**g**) and apo ScleRS (Alpha-Fold DB: AF-P09436-F1, colored in pink) (**h**) indicated the conformational changes in the C-terminal domains and ABD of ScleRS upon tRNA^{Ile} binding.

structure. Compared to this predicted structure, in addition to the ABD and C-ter A domain, the C-ter B and C domains also rotated approximately 20–30° during tRNA^{Ile} binding (Fig. 2h). Thus, in the editing state of ScleRS, the C-terminal domains must undergo conformational movements to clamp tRNA^{Ile} together with the aminoacylation main body. However, additional experimental data are needed to further elucidate the potential domain movements, which are currently inferred primarily from the AI-predicted structure.

The mechanism for recognition of anticodon A35/U36

The backbone of the anticodon loop sits in a positively charged cavity (Fig. 3a) and forms extensive electrostatic and H-bonding interactions with multiple residues from the C-ter A, C-ter C and ABD domains (Supplementary Fig. 7), including Asn660, whose mutation leads to the

resistance of eukaryotic IleRS to the natural product inhibitor reveromycin A^{11,30}. Notably, Arg739 from the ABD plays a key role in anticodon loop binding by inserting its side chain into the cavity between C32, U33, A35 and U36. It forms six H-bonds with these bases, three of which contribute to the direct recognition of A35 and U36, the second and third nucleotides of the anticodon triplet (Fig. 3b). Arg739 is well-conserved among all aligned eukaryotic IleRSs (Supplementary Fig. 6), and its mutation to alanine in ScleRS caused complete failure of tRNA^{Ile} isoleucylation in vitro (Fig. 2d). The important role of Arg739 in isoleucylation may explain the clinical finding that the compound heterozygous variants of human cytoplasmic IleRS, R739C (corresponding to Arg739 in ScleRS) and F556S (a mutation in the CD that impairs enzyme function), caused growth delay, hepatic dysfunction, and neurodevelopmental disabilities³¹. Notably, R739A and

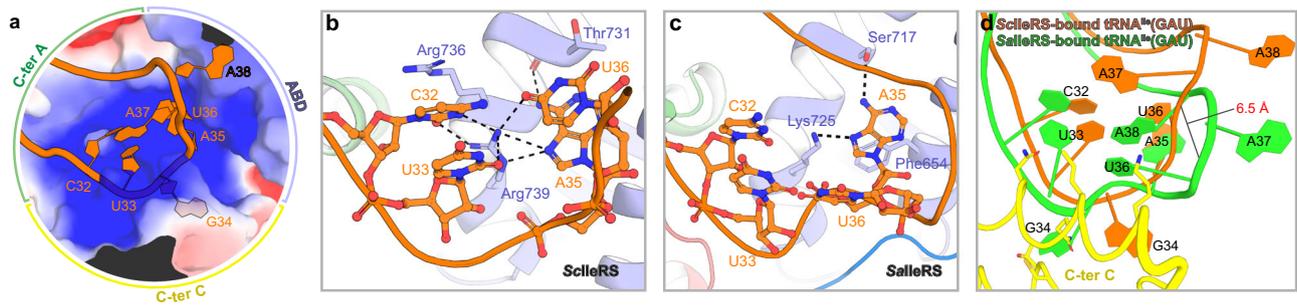


Fig. 3 | Recognition of A35/U36 by the ABD of SclleRS. **a** Electrostatic surface potential of the positively charged cavity formed by the ABD, C-ter A, and C-ter C domains for binding the anticodon loop of tRNA^{Ile}(GAU). **b** Base-specific interactions between the anticodon loop of tRNA^{Ile}(GAU) and the ABD of SclleRS. **c** As shown in the SallerRS-tRNA^{Ile}(GAU)-mupirocin complex structure (PDB ID: 1FFY), the

ABD of SallerRS only forms base-specific interactions with A35 of tRNA^{Ile}(GAU). **d** Different conformations of the anticodon loop between SallerRS-bound and SclleRS-bound tRNA^{Ile}(GAU) molecules. The nucleotides U33, G34 and U36 in the SallerRS-bound tRNA^{Ile}(GAU) clash with the C-ter C domain of SclleRS.

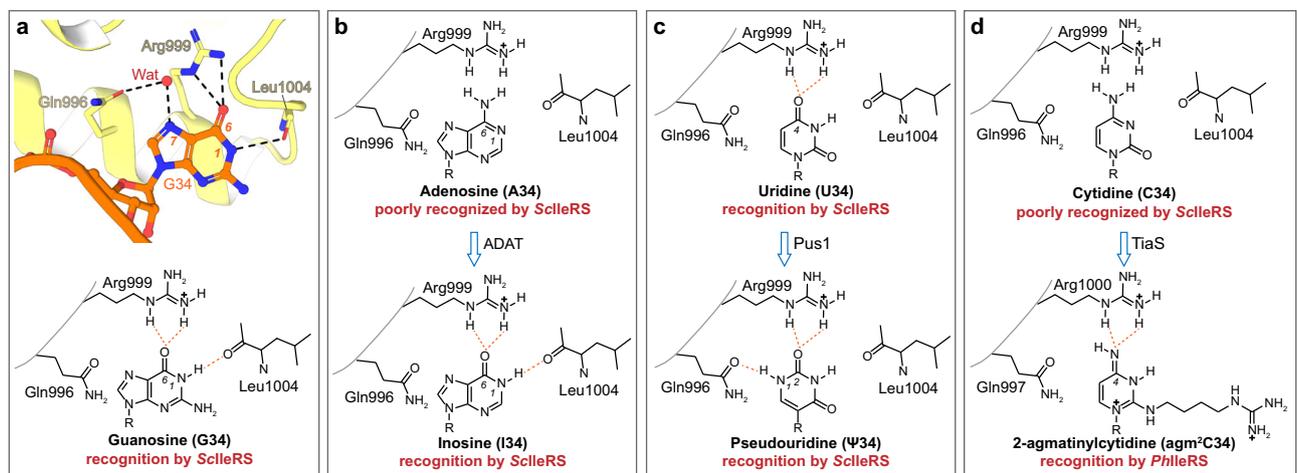


Fig. 4 | Structural mechanism of N34 recognition by the C-ter C domain of eukaryotic/archaeal-type IleRS. **a** The binding of G34 to the C-ter C domain of SclleRS. The top panel shows a zoomed-in view of the G34 binding site of the SclleRS-tRNA^{Ile}(GAU)-L-Ile complex. Direct and water-mediated H-bonding interactions are shown as black dashed lines. The bottom panel is the 2D presentation of G34-residues interactions, and the direct H-bonds are shown as orange dashed lines. The water-mediated H-bonding interactions have been omitted in the 2D

presentation. **b** 2D presentations of the modeling of A34 and I34 at the N34 binding site of SclleRS. A34 is unable to interact with the N34 binding residues due to the lack of appropriate H-bond acceptors or donors, so it is poorly recognized by SclleRS. In contrast, after deamination, I34 can bind to these residues like G34, and be well recognized by SclleRS. **c** Modeling of U34 and Ψ34 at the N34 binding site of SclleRS. **d** Modeling of C34 and agm²C34 at the N34 binding sites of SclleRS and PhlIleRS, respectively.

other site-directed mutations of SclleRS discussed later all exhibited activity comparable to that of wild-type SclleRS in the tRNA-independent pre-transfer editing assay (Fig. 2c), indicating that their inactivity in tRNA^{Ile} isoleucylation is due to deficiencies in appropriate tRNA^{Ile} binding for L-Ile transfer.

The Arg739-based A35/U36 recognition mechanism is specific to eukaryotic IleRS but not to bacterial IleRS, which relies on both the unique binding conformations of the tRNA^{Ile} anticodon loop and the anticodon loop-binding residues of IleRS. In the SclleRS-tRNA^{Ile}(GAU) complex, C32 stacks with U33, and they are both buried in a cavity of SclleRS. Arg736 contributes to stabilizing this conformation by forming a cation- π interaction with the cytosine ring of C32 (Fig. 3b), and its mutation to alanine abolished the tRNA^{Ile} isoleucylation activity of SclleRS (Fig. 2d). In contrast, C32 and U33 are partially exposed to the solvent in the SallerRS-tRNA^{Ile}(GAU)-mupirocin complex (PDB ID: 1FFY)⁵ (Fig. 3c, d). Moreover, although A35 stacks with U36, and they both point to the ABD in SclleRS-tRNA^{Ile}(GAU) complex (Fig. 3b), U36 of SallerRS-bound tRNA^{Ile}(GAU) is directed inside the anticodon loop (Fig. 3c, d). Thus, unlike that of SclleRS, the ABD of SallerRS only forms

base-specific interactions with A35 but not with C32, U33 or U36 (Fig. 3c). The distinct conformations of U36 in tRNA^{Ile}(GAU) bound to SclleRS and SallerRS propagate to the following nucleotides A37 and A38 (Fig. 3d). When the ABDs of the two IleRSs were superimposed, the anticodon loop backbones of the two tRNA^{Ile}(GAU) molecules shifted by up to 6.5 Å (Fig. 3d).

Notably, the ABD of eukaryotic SclleRS makes considerably more contacts with the anticodon loop than that of bacterial SallerRS does, but it does not recognize the nucleotide N34 to discriminate tRNA^{Ile} from tRNA^{Met}; this task must be performed with additional C-terminal domains.

A conserved arginine recognizes a common carbonyl of N34s
The nucleotide G34 of tRNA^{Ile}(GAU) flips out from the anticodon loop to form multiple polar interactions with the C-ter C domain (Fig. 4a): the N1 atom of G34 H-bonds with the backbone oxygen of Leu1004; the 6-carbonyl group forms two H-bonds with Arg999; and N7 forms a water-mediated H-bond with Gln996. While the L-Ile tRNA transfer is the rate-limiting step in tRNA^{Ile} isoleucylation compared to L-Ile

activation³², ScleRS bearing the Q996A mutation catalyzed tRNA^{leu}(GAU) isoleucylation at a rate comparable to that of the wild-type ScleRS, so the Q996A mutation was unlikely to significantly affect tRNA^{leu} recognition by ScleRS. In contrast, ScleRS bearing the R999A mutation completely lost aminoacylation activity, indicating that Arg999 plays a more important role in G34 recognition (Fig. 2d).

In addition to G34, ScleRS also recognizes U34 as well as I34 and Ψ34 (the modified A34 and U34) but poorly accommodates unmodified A34 and C34⁴. When G34 was mutated in silico to A34 in the structural model of the ScleRS-tRNA^{leu}(GAU)-L-Ile complex, A34 lost direct contacts with Arg999 and Leu1004 because its chemical groups at positions 1 and 6 have H-bond acceptor-donor properties opposite to those of G34 (Fig. 4b and Supplementary Fig. 8). After the deamination of A34, product I34 has the same chemical properties as G34 at positions 1 and 6, so I34 can bind to ScleRS in a manner similar to G34 (Fig. 4b and Supplementary Fig. 8). Interestingly, the modified nucleotide I can also function as a mimic of G to pair with C in both translation and splicing of mRNA^{33,34}, highlighting a similar recognition mechanism of I by both protein and RNA. U34 could H-bond with Arg999 via its 4-carbonyl group, and after modification, Ψ34 may form an additional H-bond with Gln996 through its N1 atom (Fig. 4c and Supplementary Fig. 8). This new H-bond can partially explain the 40-fold increase in the activity of fully modified tRNA^{leu}(ΨAΨ) compared with the unmodified tRNA^{leu}(UAU) transcribed in vitro⁴. Notably, compared with G34 and I34, U34 and Ψ34 are smaller in size and may not interact with the backbone of Leu1004 (Supplementary Fig. 8).

Thus, it is a unified N34 recognition mechanism of ScleRS in which Arg999 utilizes its side chain as the H-bond donor to form specific H-bonds with the carbonyl group (H-bond acceptor) of various N34s of tRNA^{leu}. In contrast, Arg999 cannot form this critical interaction with nucleotide C34 because C34 lacks an H-bond acceptor at the corresponding position (Fig. 4d and Supplementary Fig. 8), thereby preventing tRNA^{Met} from mis-isoleucylation by ScleRS. Thus, our structure highlights that the carbonyl group of N34 acts as a positive determinant of eukaryotic tRNA^{leu}. Consistently, its reader, Arg999, is conserved among eukaryotic IleRSs we aligned, except in *Drosophila melanogaster* IleRS, where it is replaced with a similar lysine residue (Supplementary Fig. 6). Mutation of Arg999 to lysine in ScleRS resulted in approximately a 70% reduction in aminoacylation activity (Fig. 2d), which is consistent with the fact that lysine is able to form only one H-bond with N34, whereas arginine can form two H-bonds. Archaeal IleRS has C-terminal domains similar to those of eukaryotic IleRS. In addition to G34, archaeal IleRS can also charge the tRNA^{leu} isoacceptor containing agm²C34, a modified C34 (Supplementary Fig. 1)⁹. According to structural modeling, *Pyrococcus horikoshii* IleRS (PhIleRS) can also use the arginine to form H-bonds with the 4-imine group (also an H-bond acceptor) of agm²C34 (Fig. 4d and Supplementary Fig. 8), supporting the unified N34 recognition mechanism of eukaryotic/archaeal-type IleRS.

The C-ter A and C domains are structurally similar (Supplementary Fig. 3). The C-ter A domain also has an arginine residue (Arg838) at the position corresponding to Arg999 of the C-ter C domain (Supplementary Fig. 9). Although Arg838 does not directly interact with tRNA^{leu}, both Arg838 and Arg999 may contribute to structural stability by interacting with the backbones of nearby loops (Supplementary Fig. 9). Mutations of Arg838 to lysine or alanine caused approximately a half-reduction or complete failure in tRNA^{leu} isoleucylation by ScleRS, respectively (Fig. 2d). We propose that the C-ter C domain was a duplication of the C-ter A domain during the evolution of eukaryotic/archaeal IleRS, and it was retained because its conserved arginine happened to provide an effective way, alternative to that of the ZBD in bacterial IleRS, to recognize N34. Considering the facts that 1) the overall structural similarity between the C-ter A and C domains is more significant than that between the C-ter A domain and its corresponding

C-terminal junction domain in bacterial IleRS, and that 2) the important arginine is conserved only in the C-ter A and C domains but not in the C-terminal junction domain (Supplementary Figs. 3 and 9), the acquisition of the C-ter C domain likely occurred later than the separation of bacterial and eukaryotic/archaeal IleRSs.

The ScleRS-tRNA^{leu}(CAU) complex in non-reactive state

Consistent with previous reports⁸, the unmodified tRNA^{leu}(CAU) (Fig. 1b) was unable to be isoleucylated by ScleRS and was instead methionylated by MetRS, suggesting that it functions like a tRNA^{Met} in the term of aminoacylation specificity (Fig. 5a). However, a C34G mutation was sufficient to restore the isoleucylation of tRNA^{leu}(CAU) (Fig. 5a). EMSA showed that tRNA^{leu}(CAU) could still interact with ScleRS, although its interaction is weaker than that of tRNA^{leu}(GAU) (Fig. 5b). To understand how the Arg999-N34 interaction determines whether a tRNA should be isoleucylated by ScleRS, we solved a cocrystal structure of the ScleRS-tRNA^{leu}(CAU)-L-Ile ternary complex at 2.83 Å resolution with $R/R_{free} = 23.8\%/27.8\%$ (Fig. 5c, Supplementary Fig. 2 and Supplementary Table 1). tRNA^{leu}(CAU) is still located between the aminoacylation main body and the C-terminal domains of ScleRS, but it has fewer interactions with ScleRS. The interface area of 513 Å² is much smaller than that between tRNA^{leu}(GAU) and ScleRS (2,362 Å²) as measured by program PISA³⁵. Notably, the entire anticodon loop of tRNA^{leu}(CAU), including C34, loses interactions with ScleRS and becomes too dynamic to be traced in the density map. To our surprise, the acceptor stem of tRNA^{leu}(CAU) points to neither the editing site nor the aminoacylation cavity, but binds to the backside of the ED (Fig. 5d). In this non-reactive conformation, only the nucleotides G2, C72 and C73 of the tRNA^{leu}(CAU) acceptor stem H-bond with the residues Gln393 and Asn201 of ScleRS ED, and the 3' CCA end of tRNA^{leu}(CAU) is exposed to the solution and invisible.

Two ScleRS-tRNA complex structures were aligned based on the CD of ScleRS (Fig. 5e, f), and a difference of approximately 25° was observed between tRNA^{leu}(CAU) and tRNA^{leu}(GAU) (Fig. 5e). The C-ter A and C-ter B domains of tRNA^{leu}(CAU)-bound ScleRS exhibited conformations quite similar to those of tRNA-free ScleRS (PDB ID: 7D5C and AlphaFold DB: AF-PO9436-F1) but not to those of tRNA^{leu}(GAU)-bound ScleRS (Fig. 5f). In contrast, the conformation of the distal C-ter C domain of tRNA^{leu}(CAU)-bound ScleRS is dramatically different from those of both tRNA-free and tRNA^{leu}(GAU)-bound ScleRS (Fig. 5f). Thus, the dynamic anticodon loop of tRNA^{leu}(CAU) did not attract but probably even kicked away the C-ter C domain in the ScleRS-tRNA^{leu}(CAU) complex. As a result, the C-ter C domain of tRNA^{leu}(CAU)-bound ScleRS adopts a more open conformation, and it must rotate approximately 60° and 40° to align with the C-ter C domain of tRNA^{leu}(GAU)-bound and tRNA-free ScleRS, respectively (Fig. 5f).

Based on the above structural observations and the biochemical results, we propose a possible route for the discriminative aminoacylation of tRNA^{leu} (Supplementary Fig. 10): both tRNA^{leu} and tRNA^{Met} could initially dock to ScleRS based on their rough shape and charge complementarity; ScleRS-tRNA^{leu} binding is then strengthened by the formation of multiple stacking and bonding interactions between the anticodon triplets of tRNA^{leu} and ScleRS, in which the H-bonds between the N34 of tRNA^{leu} and Arg999 of ScleRS contribute an indispensable part of the binding energy; in the meantime, mutual structural adaptation between tRNA^{leu} and ScleRS, including the rotation of tRNA^{leu} along with the C-terminal domains of ScleRS, stabilizes the complex in an editing conformation; tRNA^{leu} can then move its acceptor arm to the aminoacylation site to get charged and subsequently move it back to the editing site for proofreading, followed by its release from ScleRS for protein translation. Because C34 cannot contribute to H-bonding with Arg999, the binding of tRNA^{Met} to ScleRS will be suspended at the first step, after which tRNA^{Met} will dissociates from ScleRS without being mis-isoleucylated.

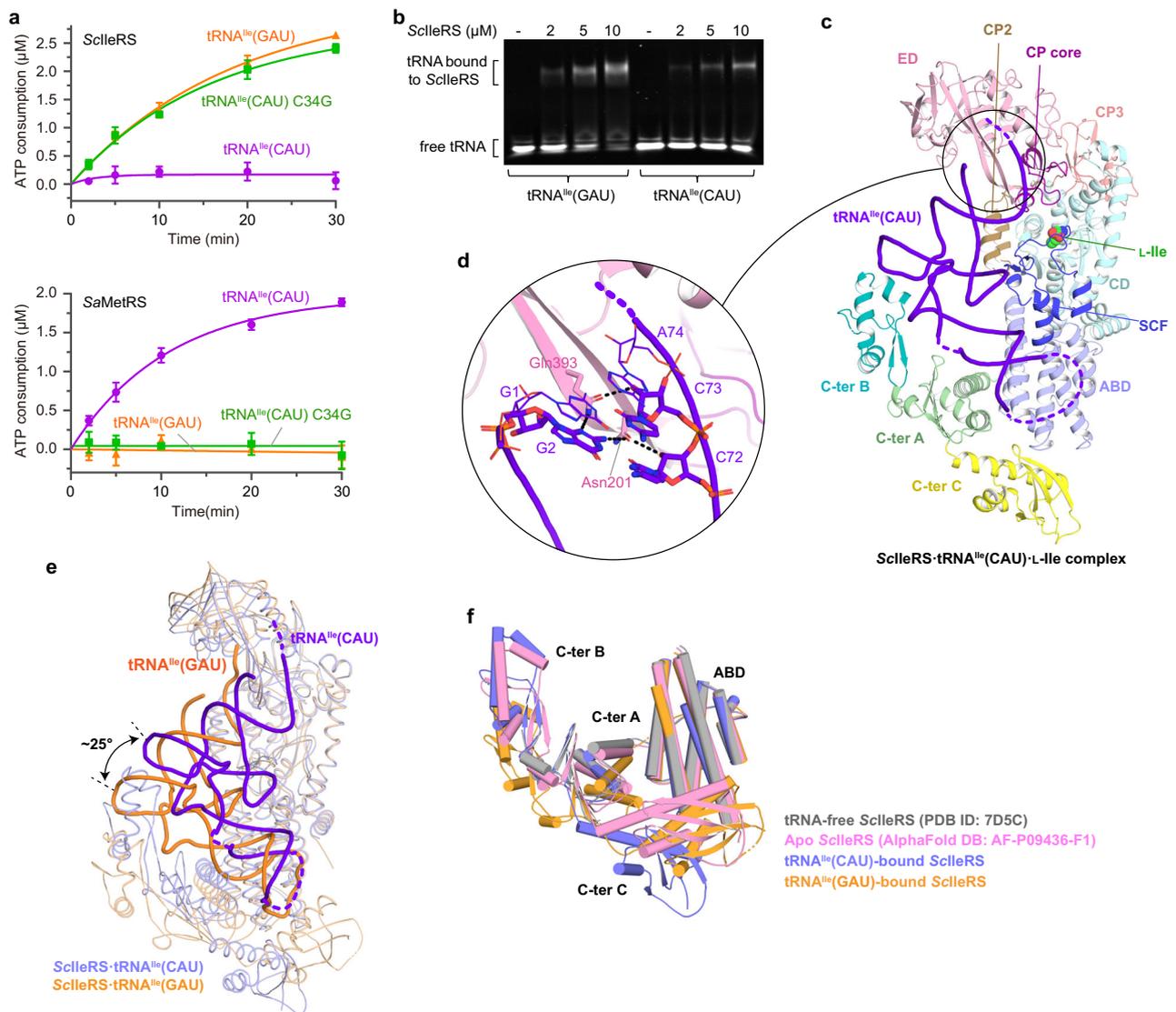


Fig. 5 | The ScleRS-tRNA^{lle}(CAU)-L-Ile complex is stalled in a non-reactive state due to the lack of key N34 interactions. **a** The aminoacylation activities of ScleRS and SaMetRS against in vitro transcribed tRNA^{lle}(GAU), tRNA^{lle}(CAU) and tRNA^{lle}(CAU) C34G variant. Data are presented as means \pm SD ($n = 3$ independent experiments). **b** EMSA revealed that in vitro transcribed tRNA^{lle}(CAU) can still form a complex with ScleRS, although it is slightly weaker than in vitro transcribed tRNA^{lle}(GAU). Similar results were observed in two independent experiments. **c** The overall structure of the ScleRS-tRNA^{lle}(CAU)-L-Ile complex. The anticodon loop of tRNA^{lle}(CAU) is dynamic because of the lack of interactions with ScleRS, and the

acceptor stem of tRNA^{lle}(CAU) binds to the back of the ED. **d** Binding of the acceptor stem of tRNA^{lle}(CAU) to the back of ScleRS ED. The 3' CCA end was invisible in the electronic density map. **e** Structural comparison between the reactive and non-reactive ScleRS-tRNA complexes revealed an approximately 25° rotation between tRNA^{lle}(GAU) and tRNA^{lle}(CAU). **f** The C-terminal domains of tRNA^{lle}(CAU)-bound ScleRS adopt a conformation generally similar to that of ScleRS in the tRNA-free state except for the distal C-ter C domain, but not to that of tRNA^{lle}(GAU)-bound ScleRS.

Discussion

In this study, the cocrystal structures showed that the complex of ScleRS with the cognate tRNA^{lle}(GAU) adopts an editing conformation, while the complex of ScleRS with the unmodified tRNA^{lle}(CAU) remains in a non-reactive state. Many structural studies have indicated that class Ia AARSs with ED (IleRS, LeuRS and ValRS) bind their substrate tRNAs mostly stably in the editing state^{5,18,19}. The ED is known to ensure the fidelity of tRNA charging through amino acid proofreading. This fidelity check occurs prior to the release of aminoacyl-tRNA and is performed using the ED's amino acid pocket and catalytic activity. Here, the fact that only the cognate tRNA^{lle}(GAU) but not the unmodified tRNA^{lle}(CAU) can bind to the ED of ScleRS at the correct editing conformation suggests that the ED likely also facilitates the fidelity check of tRNA. This tRNA check is performed upon tRNA entry using the ED in conjunction with other tRNA binding domains. Thus, the ED

likely plays fidelity check roles for both amino acids and tRNAs, but in two distinctive ways.

The accurate recognition of N34 by IleRS plays the most important role in the discrimination of tRNA^{lle} from tRNA^{Met} and subsequently maintains the fidelity of protein translation. Bacterial and eukaryotic/archaeal IleRSs recruit kingdom-specific domains to recognize cognate N34s. Bacterial SallERS was shown to recognize the guanine ring of G34 of tRNA^{lle}(GAU) by stacking and H-bonding interactions with its Trp890 and Arg888 from the ZBD (PDB ID: 1FFY)⁵ (Supplementary Fig. 11). Interestingly, although both the ZBD and the C-ter C domain H-bond with G34 via an arginine residue, the ZBD utilizes the main chain carbonyl group of arginine (Arg888 in SallERS) as an H-bond acceptor to interact with N1 and 2-NH₂ of G34, while the C-ter C domain utilizes the side chain guanidyl group of arginine (Arg999 in ScleRS) as an H-bond donor to interact with the

6-carbonyl group of G34. Moreover, the orientations of G34 are completely opposite in SclleRS- and ScleRS-bound tRNA^{le} (Fig. 3d). Therefore, the recognition mechanisms of G34 by bacterial and eukaryotic/archaeal IleRSs are completely independent of each other in both the recognition domain and the binding mode.

Eukaryotic/archaeal IleRS utilizes a unified mechanism to recognize all the N34s of their cognate tRNA^{le} substrates (Fig. 4 and Supplementary Fig. 8). Mupirocin, a natural product that selectively inhibits bacterial IleRS, is widely used to treat skin infections. However, some bacteria acquire *mupA* or *mupB* genes from the environment, which express mupirocin-resistant IleRS (MupA or MupB), resulting in the ineffectiveness of mupirocin in clinic^{36,37}. In addition, some bacteria, such as the mupirocin-producing *Pseudomonas fluorescens*, contain an endogenous mupirocin-resistant IleRS (IleRS2)^{38,39}. These mupirocin-resistant IleRSs contain eukaryotic/archaeal-type C-terminal domains⁴⁰, and can isoleucylate bacterial tRNA^{le} with L34 to decode the AUA codon on mRNA⁴¹. Structural modeling suggested that the C-ter C domain of PflIleRS2 also employs a conserved arginine to H-bond with the 4-imine group of L34 of bacterial tRNA^{le} (Supplementary Fig. 8). Thus, the arginine-mediated N34 recognition mechanism of eukaryotic/archaeal-type IleRS could also apply to the recognition of N34 modifications of bacterial tRNA^{le}, but bacterial IleRS is unable to cross-charge eukaryotic tRNA^{le}¹², suggesting that eukaryotic/archaeal-type IleRS likely has developed a more robust way for N34 discrimination than its bacterial counterpart.

The unusual non-reactive complex of ScleRS with C34-unmodified tRNA^{le}(CAU) highlights the important roles of N34 modifications in IleRS-tRNA recognition. Bacterial, archaeal and eukaryotic tRNA^{le} molecules have developed their own modifications on N34 (such as Ψ34, I34, L34 and agm²C34)^{4,8,9}. Thus, the N34 recognition domains and key residues on IleRS as well as the anticodon loop conformations and N34 modifications of tRNA^{le} are different between bacteria, archaea and eukaryotes, presenting an example of coevolution of AARS-tRNA pairs and also providing a valuable opportunity for drugging aminoacylation with lineage specificity.

Methods

Protein preparation

The wild-type ScleRS (UniProtKB ID: P09436) was expressed and purified as described¹¹. Briefly, the DNA sequence encoding ScleRS was cloned into the pET20b(+) plasmid (Novagen) with a C-terminal hexahistidine tag. The *E. coli* BL21(DE3) cells transformed with IleRS-pET20b(+) were grown in Luria-Bertani (LB) medium supplemented with 0.1 mg/mL ampicillin at 37 °C until the OD₆₀₀ reached approximately 0.6, and then 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) was added to induce protein overexpression at 20 °C for 16 h. The cells were harvested by centrifugation, resuspended in washing buffer (400 mM NaCl, 50 mM Tris-HCl pH 8.0, 10% *v/v* glycerol, 5 mM β-ME and 10 mM imidazole) and lysed by sonication. The lysate was centrifuged at 4000 × g for 30 min to remove cell debris. The supernatant was loaded onto a Ni-NTA column (Qiagen) and washed with 20 column volumes of washing buffer. The target protein was eluted with 5 column volumes of elution buffer (400 mM NaCl, 50 mM Tris-HCl pH 8.0, 10% *v/v* glycerol, 5 mM β-ME and 100 mM imidazole). The elution was concentrated to 2 mL and injected into a HiLoad 16/60 Superdex 200 pg (GE healthcare) column. The peak fractions eluted using gel filtration buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM β-ME and 5% *v/v* glycerol) were collected and concentrated to ~30 mg/mL. The plasmids for overexpressing ScleRS mutants were constructed by amplifying the whole plasmid using PCR. For each mutation, a pair of primers that partially overlap and both contain the targeted mutation were used (Supplementary Table 2). The expression and purification of all ScleRS variants were performed in the same manner as the wild-type protein.

The DNA sequence encoding SaMetRS (Uniprot ID: V7IMS7) was inserted into pET15b, and full-length SaMetRS protein was overexpressed in BL21(DE3) and purified by Ni-NTA column (Qiagen) and HiLoad 16/60 Superdex 200 pg column (GE Healthcare). The detailed methods for producing SaMetRS were described previously⁴².

In vitro transcription and purification of tRNA

The *E. coli* tRNAs were produced using in vitro T7 RNA polymerase transcription assay as described¹¹. The DNA templates for tRNA^{le}(GAU) transcription were generated by PCR using Primer1 (5'-**TAA-TACGACTCACTATAGGGCTTGTAGCTCAGGTGGTTAGAGCGCACCC-CTGATAAG-3'**) and Primer2 (5'-TGGTGGGCTGAGTGGACTTGAAC-CACCGACCTCACCCATTACAGGGGTGGCCTCTAAC-3'). This two primers cover the full-length of tRNA^{le}(GAU) gene and are partially complementary to each other (the underlined nucleotides), and Primer1 contains T7 promoter sequence (nucleotides in bold). The wild-type base pair A1-U72 of *E. coli* tRNA^{le}(GAU) was replaced by the G1-C72 pair to increase transcription³. The PCR product was further amplified by the second round of PCR using primer3 (5'-TAATACGACTCACTATAGGGCTTGT-3') and primer4 (5'-**TGGTGGGCTGAGTGGACTT-GAAC-3'**). The first two nucleotides in italics and bold at the 5' end of primer4 were methylated at their 2'-hydroxyl groups to reduce the non-templated addition of nucleotides to the 3' end of tRNA by T7 RNA polymerase. The DNA templates for tRNA^{le}(CAU) and mutants of tRNA^{le}(CAU) and tRNA^{le}(GAU) are prepared by the same way, and the primers used are listed in Supplementary Table 3.

The product of second PCR was then used as the DNA template for in vitro transcription assay without additional purification. In a 15 mL centrifugal tube, 2 mL of the PCR product was mixed with 2 mL of 5 × transcription buffer (1 M Tris pH 8.0, 10 mM spermidine and 50 mM DTT), 1 mL of 40 mM NTPs (each), 0.2 mL of 1 M MgCl₂, 0.2 mL of 10 mg/mL T7 polymerase and 3.6 mL of DECP water, and incubated at 37 °C for 3 to 4 h. The transcripts were purified using 12% polyacrylamide gel electrophoresis supplemented with 8 M urea, extracted from gels by 0.5 M ammonium acetate, and precipitated by ethanol at -20 °C overnight. The tRNA pellets were collected by centrifugation, washed by 70% *v/v* ethanol, and redissolved in a buffer consisting of 20 mM Tris pH 8.0 and 1 mM EDTA to 1 mg/mL. The redissolved tRNA was heated at 65 °C for 5 min, and then refolded by slowly cooling to room temperature after the addition of 10 mM MgCl₂. The refolded tRNA was concentrated to ~10 mg/mL using a 3 kDa Ultra-4 centrifugal filter device (Millipore), aliquoted and stored at -80 °C for further use.

Overexpression and purification of tRNA

The *E. coli* tRNA^{le}(GAU) gene with the native base pair A1-U72 substituted with G1-C72 was inserted between the T7 promoter and terminator of pET29b(+) by homologous recombination. The transformed *E. coli* BL21(DE3) cells were cultured in LB medium supplemented with 50 μg/mL kanamycin until the OD₆₀₀ reached approximately 0.6, and then 1 mM IPTG was added to induce the overexpression of tRNA^{le}(GAU) at 30 °C for 16 h. The tRNA transcript was extracted from the cell pellets using RNAiso Plus (Cat. No. 9109, Takara) and chloroform, and precipitated from aqueous fractions by isopropanol. The tRNA pellets collected by centrifugation were washed with 70% *v/v* ethanol and redissolved in a buffer containing 20 mM Tris pH 8.0 and 10 mM MgCl₂. The sample was loaded onto a HiTrap Q XL (GE healthcare) column and the elution fractions between 0.55 and 0.70 M NaCl were collected. Finally, the tRNA was concentrated to ~20 mg/mL and stored in a buffer containing 10 mM HEPES pH 7.5 and 10 mM MgCl₂.

Crystallography

The sitting-drop vapour-diffusion method was employed to crystallize the ScleRS-tRNA complex. The full-length ScleRS (10 mg/mL) was preincubated with tRNA^{le}(GAU) or tRNA^{le}(CAU) (transcribed in vitro,

2.5 mg/mL), together with 5 mM L-Ile, at room temperature for 30 min. Each drop containing 1 μ L of protein-tRNA mixture, 0.5 μ L of reservoir solution and 0.5 μ L of seed stocked in reservoir solution was equilibrated against 100 μ L of reservoir solution at 8 °C for 3–7 days to allow crystals to grow. For ScIleRS-tRNA^{Ile}(GAU) complex, the reservoir solution contains 0.2 M ammonium sulfate, 0.1 M BIS-TRIS pH 5.5, 25% PEG3350 and 0.06 M sodium citrate. For ScIleRS-tRNA^{Ile}(CAU) complex, the reservoir solution contains 2% Tacsimate pH 6.0, 0.1 M BIS-TRIS pH 6.5 and 20% PEG3350. Large crystals were immersed in a reservoir solution supplemented with 20% ethylene glycol for a few seconds and then flash frozen in liquid nitrogen. The diffraction data were collected using a single crystal at 100 K with a wavelength of 0.979 Å at the BL19U1 beamline at National Facility for Protein Sciences Shanghai (NFPS) and Shanghai Synchrotron Radiation Facility (SSRF) and were indexed, integrated and scaled using XDS⁴³ and Aimless⁴⁴. The structure was solved by molecular replacement using the ScIleRS structure (PDB ID: 7D5C)¹¹ as the search model in the program Molrep⁴⁵. Iterative refinements of the structure model were carried out using Coot⁴⁶ and Refmac5⁴⁷. The stereochemical quality of the final model was assessed using MolProbity⁴⁸. The statistics of the data collection and structural refinement are listed in Supplementary Table 1. Final Ramachandran statistics were as follows: 94.4% favored, 5.5% allowed and 0.1% outliers for the ScIleRS-tRNA^{Ile}(GAU)-L-Ile complex; 93.6% favored, 6.3% allowed and 0.1% outliers for the ScIleRS-tRNA^{Ile}(CAU)-L-Ile complex. The coordinate and structural factors of the ScIleRS-tRNA^{Ile}(GAU)-L-Ile and ScIleRS-tRNA^{Ile}(CAU)-L-Ile complex have been deposited in the Protein Data Bank (PDB) under the accession code 8WND and 8ZIP respectively.

ATP consumption assay

ATP consumption assay was employed to evaluate the aminoacylation activities of ScIleRS and its variants on tRNA^{Ile}. The 60 μ L of reactions contained 40 nM ScIleRS (wild type or variants), 200 μ M ATP, 1 mM L-isoleucine, 1 mg/mL *E. coli* tRNA^{Ile}(GAU) (overexpressed in *E. coli*) in the reaction buffer (30 mM HEPES pH 7.5, 150 mM NaCl, 30 mM KCl, 40 mM MgCl₂, 1 mM DTT and 0.1% BSA). The reaction was incubated at room temperature. Aliquots of 5 μ L at various time points (2, 5, 10, 20 and 30 min) were transferred to a 384 well plate and mixed with 5 μ L of Kinase-Glo[®] Max Reagent (Cat. No. V6071, Promega) to stop the reaction. After 20 min, the luminescence (L) which reflects the concentration of the remaining ATP, was read on a Synergy HI microplate reader (BioTek). The reactions without the addition of tRNA were used as controls (Lc). The ATP consumption (μ M) = 200 × (1-L/Lc). Each reaction was repeated three times, and the results are expressed as the means ± SD (n = 3). Statistical analyzes were performed using GraphPad Prism 7 software, and a one-phase association equation was used to fit the time response curves (ATP consumption vs reaction time).

To evaluate the aminoacylation activities and specificities of in vitro tRNA transcripts, 60 μ L of reactions containing 50 nM ScIleRS or SaMetRS, 4 μ M ATP, 1 mM L-isoleucine or L-methionine and 1 mg/mL in vitro transcribed tRNA were incubated at room temperature. Aliquots of 5 μ L at various time points (2, 5, 10, 20 and 30 min) were transferred to a 384 well plate and mixed with 5 μ L of Kinase-Glo[®] Reagent (Cat. No. V6711, Promega) to stop the reaction. After 10 min, the luminescence (L) was read and the ATP consumption was calculated. The time response curves (ATP consumption vs reaction time) were fitted.

tRNA-independent Pre-transfer editing assay

Assays were performed in 80 μ L of reaction mixtures consisting of 80 nM wild-type ScIleRS or its mutants, 6 mM L-cysteine, 250 μ M ATP, 50 μ g/mL PPIase, 30 mM HEPES pH 7.5, 150 mM NaCl, 30 mM KCl, 40 mM MgCl₂ and 1 mM DTT. After incubation at room temperature for 30 min, 20 μ L of malachite green reagent (2.45 M sulfuric acid, 0.1%

malachite green, 1.5% ammonium molybdate tetrahydrate and 0.2% tween-20) was added to the mixtures. After incubation for 10 min, absorbance (A) was measured at 620 nm. The reactions without the addition of ScIleRS were used as controls. The absorbance differences between the wells with ScIleRS and those without ScIleRS reflected the tRNA-independent pre-transfer editing activity of ScIleRS. The tRNA-independent pre-transfer editing activity of the wild-type ScIleRS was normalized to 100%. Each reaction was repeated three times, and the results are expressed as the means ± SD (n = 3).

Electrophoresis mobility shift assay (EMSA)

The 20 μ L reaction mixtures containing 1 μ M in vitro transcript of tRNA^{Ile}(GAU) or tRNA^{Ile}(CAU) and ScIleRS at different concentrations (2, 5 and 10 μ M) in binding buffer (30 mM sodium cacodylate pH 6.5, 20 mM MgCl₂, 150 mM NaCl, 2 mM DTT and 10% v/v glycerol) were incubated at 4 °C for 10 min. The samples were loaded to 5% native polyacrylamide gels and electrophoresed at a voltage of 80 V for 2 h on a ice-bath. The gel was stained with Gel-Red nucleic acid dye (BBI Life Sciences).

Data analysis and figure preparation

Multiple protein sequence alignment and conservation score calculation were performed using Clustal Omega program⁴⁹ and Jalview program⁵⁰. All data were analyzed using GraphPad Prism 8.0 software and are expressed as the means ± SD (n = 3). All protein structure figures were prepared using PyMOL (PyMOL v.2.5.0).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available from the corresponding authors upon request. The coordinates and structural factors of ScIleRS-tRNA^{Ile}(GAU)-L-Ile and ScIleRS-tRNA^{Ile}(CAU)-L-Ile complexes have been deposited in the Protein Data Bank (PDB) under the accession codes 8WND and 8ZIP. The structures used for molecular replacement or structural analyses are publicly available in PDB under accession codes 1FFY, 1WNZ, 4AQ7, 7D5C, 1ILE, 6LDK, 8WNF and 4QE1. The source data for Figs. 2c–f and 5a, b are provided as a Source data file. Source data are provided with this paper.

References

- Carter, C. W. Jr. Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* **62**, 715–748 (1993).
- Ibba, M. & Soll, D. Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**, 617–650 (2000).
- Nureki, O. et al. Molecular recognition of the identity-determinant set of isoleucine transfer RNA from *Escherichia coli*. *J. Mol. Biol.* **236**, 710–724 (1994).
- Senger, B., Auxilien, S., Englisch, U., Cramer, F. & Fasiolo, F. The modified wobble base inosine in yeast tRNA^{Ile} is a positive determinant for aminoacylation by isoleucyl-tRNA synthetase. *Biochemistry* **36**, 8269–8275 (1997).
- Silvian, L. F., Wang, J. & Steitz, T. A. Insights into editing from an Ile-tRNA synthetase structure with tRNA^{Ile} and mupirocin. *Science* **285**, 1074–1077 (1999).
- Pixa, G., Dirheimer, G. & Keith, G. Sequence of tRNA^{Ile}_{IAU} from brewer's yeast. *Biochem. Biophys. Res. Commun.* **119**, 905–912 (1984).
- Szweykowska-Kulinska, Z., Senger, B., Keith, G., Fasiolo, F. & Grosjean, H. Intron-dependent formation of pseudouridines in the anticodon of *Saccharomyces cerevisiae* minor tRNA^{Ile}. *EMBO J.* **13**, 4636–4644 (1994).

8. Muramatsu, T. et al. Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature* **336**, 179–181 (1988).
9. Ikeuchi, Y. et al. Agmatine-conjugated cytidine in a tRNA anticodon is essential for AUA decoding in archaea. *Nat. Chem. Biol.* **6**, 277–282 (2010).
10. Chung, S., Kim, S., Ryu, S. H., Hwang, K. Y. & Cho, Y. Structural basis for the antibiotic resistance of eukaryotic isoleucyl-tRNA synthetase. *Mol. Cells* **43**, 350–359 (2020).
11. Chen, B. et al. Inhibitory mechanism of reveromycin A at the tRNA binding site of a class I synthetase. *Nat. Commun.* **12**, 1616 (2021).
12. Shiba, K. et al. Human cytoplasmic isoleucyl-tRNA synthetase: selective divergence of the anticodon-binding domain and acquisition of a new structural unit. *Proc. Natl Acad. Sci. Usa.* **91**, 7435–7439 (1994).
13. Chung, S. et al. Regulation of BRCA1 stability through the tandem UBX domains of isoleucyl-tRNA synthetase 1. *Nat. Commun.* **13**, 6732 (2022).
14. Naganuma, M. et al. The selective tRNA aminoacylation mechanism based on a single G•U pair. *Nature* **510**, 507–511 (2014).
15. Fersht, A. R. Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry* **16**, 1025–1030 (1977).
16. Bilus, M. et al. On the mechanism and origin of isoleucyl-tRNA synthetase editing against norvaline. *J. Mol. Biol.* **431**, 1284–1297 (2019).
17. Fukunaga, R. & Yokoyama, S. Structural basis for substrate recognition by the editing domain of isoleucyl-tRNA synthetase. *J. Mol. Biol.* **359**, 901–912 (2006).
18. Fukai, S. et al. Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA^{Val} and valyl-tRNA synthetase. *Cell* **103**, 793–803 (2000).
19. Palencia, A. et al. Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. *Nat. Struct. Mol. Biol.* **19**, 677–684 (2012).
20. Nureki, O. et al. Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280**, 578–582 (1998).
21. Chen, X. et al. Structural basis for substrate and antibiotic recognition by *Helicobacter pylori* isoleucyl-tRNA synthetase. *FEBS Lett.* **598**, 521–536 (2024).
22. Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S. & Cusack, S. The crystal structure of leucyl-tRNA synthetase complexed with tRNA^{Leu} in the post-transfer-editing conformation. *Nat. Struct. Mol. Biol.* **12**, 923–930 (2005).
23. Liu, R. J. et al. Molecular basis of the multifaceted functions of human leucyl-tRNA synthetase in protein synthesis and beyond. *Nucleic Acids Res.* **48**, 4946–4959 (2020).
24. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
25. Deng, X. et al. Large conformational changes of insertion 3 in human glycyl-tRNA Synthetase (hGlyRS) during catalysis. *J. Biol. Chem.* **291**, 5740–5752 (2016).
26. Ju, Y. et al. X-shaped structure of bacterial heterotetrameric tRNA synthetase suggests cryptic prokaryote functions and a rationale for synthetase classifications. *Nucleic Acids Res.* **49**, 10106–10119 (2021).
27. Han, L. et al. The binding mode of orphan glycyl-tRNA synthetase with tRNA supports the synthetase classification and reveals large domain movements. *Sci. Adv.* **9**, eadf1027 (2023).
28. Yu, Z. et al. Structural basis of a two-step tRNA recognition mechanism for plastid glycyl-tRNA synthetase. *Nucleic Acids Res.* **51**, 4000–4011 (2023).
29. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
30. Miyamoto, Y. et al. Identification of *Saccharomyces cerevisiae* isoleucyl-tRNA synthetase as a target of the G1-specific inhibitor Reveromycin A. *J. Biol. Chem.* **277**, 28810–28814 (2002).
31. Orenstein, N. et al. Bi-allelic IARS mutations in a child with intra-uterine growth retardation, neonatal cholestasis, and mild developmental delay. *Clin. Genet.* **91**, 913–917 (2017).
32. Fersht, A. R. & Kaethner, M. M. Mechanism of aminoacylation of tRNA. Proof of the aminoacyl adenylate pathway for the isoleucyl- and tyrosyl-tRNA synthetases from *Escherichia coli* K12. *Biochemistry* **15**, 818–823 (1976).
33. Rueter, S. M., Dawson, T. R. & Emeson, R. B. Regulation of alternative splicing by RNA editing. *Nature* **399**, 75–80 (1999).
34. Mendoza, H. G. & Beal, P. A. Structural and functional effects of inosine modification in mRNA. *RNA* **30**, 512–520 (2024).
35. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
36. Hodgson, J. E. et al. Molecular characterization of the gene encoding high-level mupirocin resistance in *Staphylococcus aureus* J2870. *Antimicrob. Agents Chemother.* **38**, 1205–1208 (1994).
37. Seah, C. et al. MupB, a new high-level mupirocin resistance mechanism in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **56**, 1916–1920 (2012).
38. Yanagisawa, T. & Kawakami, M. How does *Pseudomonas fluorescens* avoid suicide from its antibiotic pseudomonon acid?: Evidence for two evolutionarily distinct isoleucyl-tRNA synthetases conferring self-defense. *J. Biol. Chem.* **278**, 25887–25894 (2003).
39. Zanki, V., Bozic, B., Mocibob, M., Ban, N. & Gruic-Sovulj, I. A pair of isoleucyl-tRNA synthetases in *Bacilli* fulfills complementary roles to keep fast translation and provide antibiotic resistance. *Protein Sci.* **31**, e4418 (2022).
40. Brkic, A. et al. Antibiotic hyper-resistance in a class I aminoacyl-tRNA synthetase with altered active site signature motif. *Nat. Commun.* **14**, 5498 (2023).
41. Suzuki, T. & Numata, T. Convergent evolution of AUA decoding in bacteria and archaea. *RNA Biol.* **11**, 1586–1596 (2014).
42. Yi, J. et al. Fragment screening and structural analyses highlight the ATP-assisted ligand binding for inhibitor discovery against type 1 methionyl-tRNA synthetase. *Nucleic Acids Res.* **50**, 4755–4768 (2022).
43. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
44. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1204–1214 (2013).
45. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 22–25 (2010).
46. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
47. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240–255 (1997).
48. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
49. Madeira, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
50. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

Acknowledgements

We thank the staff of BL19U1 beamline (<https://cstr.cn/31129.02.NFPS.BL19U1>) at the National Facility for Protein Science in Shanghai (NFPS), Shanghai Advanced Research Institute, Chinese Academy of Sciences,

for providing technical support in data collection and analysis. We also thank Prof. Xiang-Lei Yang from Scripps Research Institute for her helpful discussion on this research. This research was supported by National Natural Science Foundation of China (22177140 to H.Z. and 22207133 to B.C.), Guangdong Basic and Applied Basic Research Foundation (2023A1515012453, 2023B1515040006 and 2022B1515130008 to H.Z., 2023A1515012936 and 2021A1515110117 to B.C.), Guangdong Provincial Key Laboratory of Construction Foundation (2023B1212060022 to H.Z.) and Open Research Funds of the Affiliated Qingyuan Hospital (Qingyuan People's Hospital), Guangzhou Medical University (202301-303 to H.Z. and H.L.).

Author contributions

B.C. grew the crystals, solved the structures and performed the functional experiments. F.Y., Z.L., F.L. and S.L. contributed to the biochemical experiments. H.L. and Q.G. contributed to structural analysis and experiment design. B.C. and H.Z. wrote the manuscript. H.Z. supervised this research. All the authors have read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-55183-0>.

Correspondence and requests for materials should be addressed to Huihao Zhou.

Peer review information *Nature Communications* thanks Bernhard Kuhle, Jinwei Zhang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024