Two crystal structures reveal design for repurposing the C-Ala domain of human AlaRS

Litao Sun[a,b], Youngzee Song[a,b], David Blocquel[a,b], Xiang-Lei Yang[a,b], and Paul Schimmel[a,b,c,1]

[a]The Scripps Laboratories for tRNA Synthetase Research, The Scripps Research Institute, La Jolla, CA 92037; [b]Department of Cell and Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037; and [c]Department of Metabolism and Aging, The Scripps Florida Research Institute, Jupiter, FL 33458

Contributed by Paul Schimmel, October 24, 2016 (sent for review October 1, 2016; reviewed by Osamu Nureki and Lluís Ribas de Pouplana)

The 20 aminoacyl tRNA synthetases (aaRSs) couple each amino acid to their cognate tRNAs. During evolution, 19 aaRSs expanded by acquiring novel noncatalytic appended domains, which are absent from bacteria and many lower eukaryotes but confer extracellular and nuclear functions in higher organisms. AlaRS is the single exception, with an appended C-terminal domain (C-Ala) that is conserved from prokaryotes to humans but with a wide sequence divergence. In human cells, C-Ala is also a splice variant of AlaRS. Crystal structures of two forms of human C-Ala, and small-angle X-ray scattering of AlaRS, showed that the large sequence divergence of human C-Ala reshaped C-Ala in a way that changed the global architecture of AlaRS. This reshaping removes the role of C-Ala in prokaryotes for docking tRNA and instead repurposes it to form a dimer interface presenting a DNA-binding groove. This groove cannot form with the bacterial ortholog. Direct DNA binding by human C-Ala, but not by bacterial C-Ala, was demonstrated. Thus, instead of acquiring a novel appended domain like other human aaRSs, which engendered novel functions, a new AlaRS architecture was created by diversifying a preexisting appended domain.

appended domain | evolution | structural plasticity | DNA binding | splice variant

Mammalian aminoacyl tRNA synthetases (aaRSs) have diverse ex-translational functions that include extracellular and nuclear roles manifested in, among other functions, proangiogenesis and antiangiogenesis, immunoregulation, neurogenesis, and stress responses (1–7). These functions are considered to link aaRSs to heritable diseases (3, 4, 8). Although absent with one exception in bacteria, novel appended domains were gained in a progressive and accretive way during the evolution of eukaryotes (1, 9). These domains are dispensable for the catalytic function, but required for new nontranslational functions (1, 9). A total of 13 different appended domains have been annotated (9–11). Some, such as the WHEP [TrpRS(W), HisRS(H), GluProRS(EP)] domain, are joined to more than one tRNA synthetase, but differentiated by wide sequence divergences that are idiosyncratic to the aaRSs (1, 11–15). Most of the more than 250 recorded splice variants of human aaRSs ablate the catalytic domain but retain the noncatalytic addition (9). Whereas 19 tRNA synthetases acquired new domains during evolution, AlaRS is an exception, with a noncatalytic C-terminal domain (C-Ala) that is also present in prokaryotes (1, 10, 16).

Although C-Ala is not essential for sustaining AlaRS-dependent cell growth in bacteria, it enhances aminoacylation by providing contacts with the outside corner of the L-shaped tRNA substrate (16). It is also produced as a splice variant of human AlaRS (9). Here we used functional analysis of two crystal structures to show that human C-Ala is reshaped from docking tRNA in prokaryotes into a DNA-binding domain in humans. Thus, instead of acquiring a special appended domain, a new AlaRS architecture was created by diversifying a preexisting domain.

Results and Discussion

Human C-Ala Has No Effect on Charging Activity. The sequence of C-Ala diverged widely in the evolutionary progression to humans, and this divergence raises the possibility that C-Ala may have developed to play a different role in higher organisms and, in that respect, to be akin to the appended domains of the 19 other aaRSs. To investigate this possibility, we aligned 410 AlaRS sequences from all three kingdoms of life: eukaryotes, archae, and prokaryotes (16) (Fig. 1*A*). The alignment clearly showed the three well-characterized aminoacylation, editing, and C-Ala domains, and demonstrated that although the aminoacylation and editing domains are well conserved, C-Ala diverges widely (16) (Fig. 1*A*). Based on the crystal structure of *Archaeoglobus fulgidus* C-Ala, C-Ala consists of a helical region followed by a globular domain (17). For the various diverged C-Ala domains, we used Predictprotein (18) to predict that the helical subdomains of archaeal and bacterial C-Alas have two α-helices, whereas eukaryotes have a third, long α-helix. In contrast, the globular domains are similar across the three kingdoms (Fig. 1*B*).

The helical domain of C-Ala provides contacts for dimerization of *A. fulgidus* AlaRS (17) and for docking the outside corner of the L-shaped tRNA to the enzyme (19). Consistently, although the C-Ala segment is not essential for aminoacylation, it enhances catalytic efficiency (16, 20–22). Given the low similarity of human and bacterial or archaeal C-Ala, we compared the aminoacylation activity of full-length and C-Ala–truncated human AlaRS (*Hs* AlaRS and *Hs* AlaRS-ΔC-Ala) with their *Escherichia coli* and *A. fulgidus* orthologs. Consistent with previous studies (16, 20, 21), *Ec* or *Af* AlaRS-ΔC-Ala exhibited sharply reduced activity relative to full-length AlaRS, but in contrast, deletion of C-Ala did not

Significance

Here we present an exception that supports the rule that the 20 human tRNA synthetases acquired new architectures to expand their functions during evolution. The new features are associated with novel, appended domains that are absent in prokaryotes and retained by their many splice variants. Alanyl-tRNA synthetase (AlaRS) is the single example that has a prototypical appended domain—C-Ala—even in prokaryotes, which is spliced out in humans. X-ray structural, small-angle X-ray scattering, and functional analysis showed that human C-Ala lost its prokaryotic tRNA functional role and instead was reshaped into a nuclear DNA-binding protein. Thus, we report another paradigm for tRNA synthetase acquisition of a novel function, namely, repurposing a preexisting domain rather than addition of a new one.

Author contributions: L.S., Y.S., D.B., X.-L.Y., and P.S. designed research; L.S., Y.S., and D.B. performed research; L.S., Y.S., D.B., X.-L.Y., and P.S. analyzed data; and L.S., Y.S., and P.S. wrote the paper.

Reviewers: O.N., University of Tokyo, Graduate School of Science; and L.R.d.P., Catalan Institution for Research and Advanced Studies (ICREA) and Institute for Research in Biomedicine.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5T76 and 5T5S).

[1]To whom correspondence should be addressed. Email: schimmel@scripps.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1617316113/-/DCSupplemental.

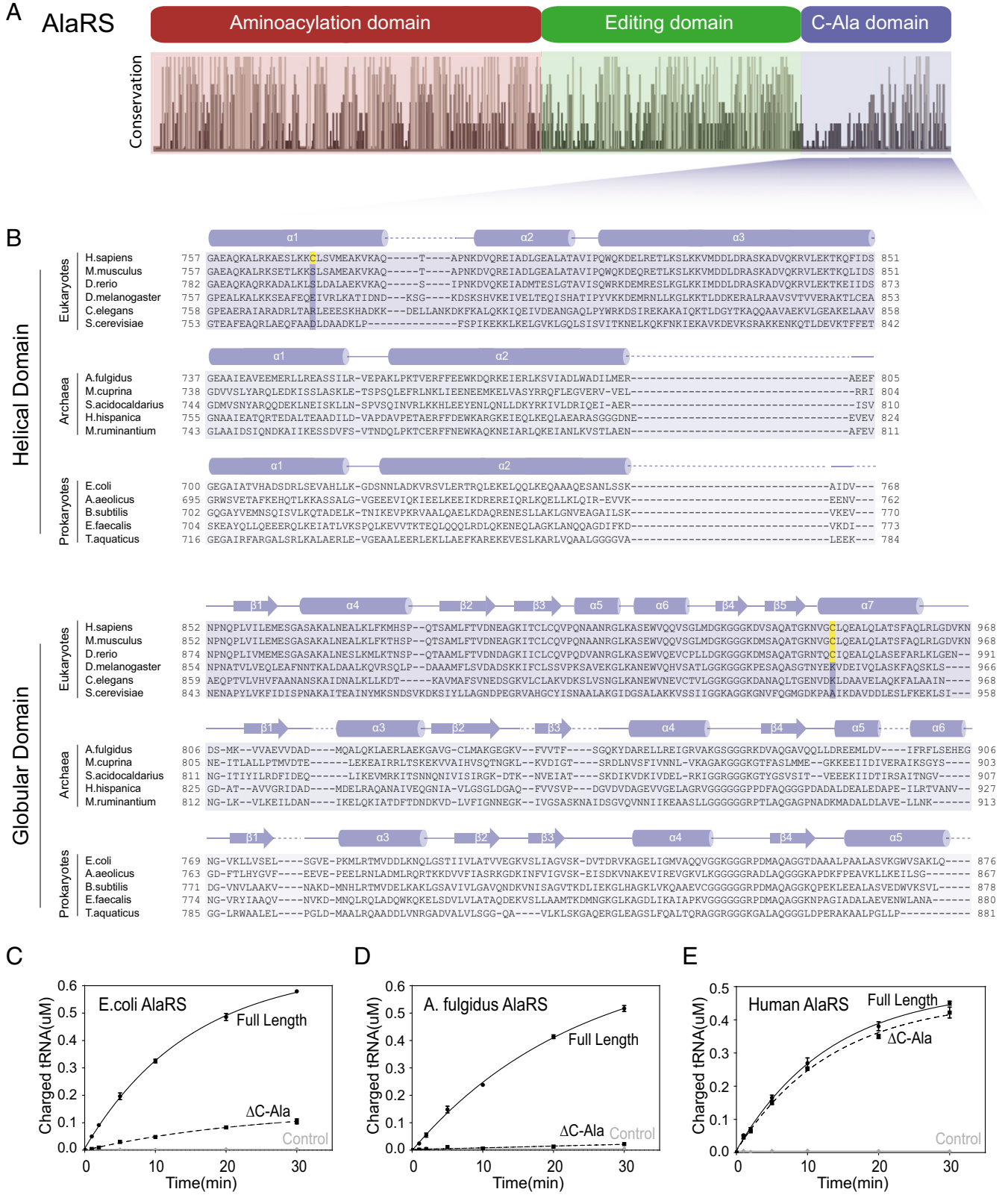


Fig. 1. Human C-Ala has no effect on the charging activity. (*A*) Conservation analysis of AlaRS sequences across bacteria, archaea, and eukaryotes showing the relative sequence identity of the 410 aligned AlaRS sequences (16). (*B*) Alignment generated using the online Clustal Omega server (28). Secondary structural elements of C-Ala are indicated above the sequences. The two cysteines (disulfide bond) are colored in yellow. (*C*, *D*, and *E*) In vitro aminoacylation assay showing that human AlaRS-ΔC-Ala has similar activity relative to human full-length AlaRS (*E*), whereas *E. coli* (*C*) or *A. fulgidus* (*D*) AlaRS-ΔC-Ala reduces the charging activity toward $tRNA^{Ala}$ compared with the corresponding full-length AlaRS. Error bars indicate SDs.

significantly affect the activity of *Hs* AlaRS (Fig. 1 *C–E*). Thus, *Hs* C-Ala is completely dispensable for aminoacylation.

Crystal Structure of Human C-Ala. Given *Hs* C-Ala's wide sequence divergence and lack of role in aminoacylation, we speculated that its structure is distinct from its prokaryote ortholog. Based on its dispensability for aminoacylation, we hypothesized that *Hs* C-Ala lacks the contact between C-Ala and the tRNA elbow region that is seen in the bacterial enzyme. Because C-Ala also contributes part of the α2 dimerization interface of bacterial AlaRS, and given the previously noted α2 dimeric quaternary structure of *Hs* AlaRS, efforts were made to understand whether *Hs* C-Ala formed a dimer. Based on secondary structure predictions and the previously solved archaeal C-Ala structure (17), we made a C-Ala construct consisting of the C-terminal 757–968 amino acids of human AlaRS. We purified the recombinant protein from the soluble fraction of the bacterial lysate and, during gel filtration, observed only a monomeric form of the recombinant human C-Ala protein in buffer containing 1 mM DTT (Fig. S1*B*). However, in addition to the monomer, a dimeric form appeared with buffer containing oxidative agents, such as 1 mM glutathione disulfide (GSSG) (Fig. S1*B*). After screening through a variety of conditions, we obtained two different crystal forms, each of which was specific to a particular condition (Fig. S1*A*). One of these crystal forms harbored the monomer and was obtained using 0.1 M Tris pH 8.5 and 25% (wt/vol) polyethylene glycol 3350, whereas the other captured a dimer using 0.2 M ammonium acetate, 0.1 M Tris pH 8.5, and 25% (wt/vol) polyethylene glycol 3350.

The structure of the monomer was determined from a selenomethionine-substituted crystal (Fig. 2*A*). This crystal (space group $P2_1$, with unit-cell parameters a = 41.645, b = 38.449, and c = 62.471Å) diffracted up to 2.0 Å with an asymmetric unit containing one molecule of C-Ala, and a refined model R_{work} factor of 21.46% and R_{free} factor of 25.76%. The dimer-containing crystal (space group $C222_1$, with unit-cell parameters a = 90.829, b = 136.141, and c = 59.077Å) had one molecule of C-Ala in the asymmetric unit, with a refined model R_{work} factor of 20.78% and and R_{free} factor of 25.45% (Fig. 2*B*). Details of the structure determination are provided in *SI Methods* and Table S1.

As expected, the monomer of *Hs* C-Ala consists of a helical subdomain and a separate globular subdomain, similar to the *A. fulgidus* C-Ala. The helical subdomain contains three α-helices, consistent with the secondary structure predictions, whereas the globular subdomain comprises a five-stranded β-sheet and four α-helices (Figs. 1*B* and 2*A*). Interestingly, in the dimeric form, a disulfide bridge was formed between the helical subdomain of one molecule and the globular subdomain of the other molecule (Fig. 2 *B* and *C* and Fig. S1*C*).

The root-mean-squared deviation (rmsd) of the Cα positions between the monomer and dimer is ∼1.5 Å for the superimposed helical subdomains and 0.6 Å for the globular subdomains (Fig. S1*D*). When the two structures were superimposed, we observed that Asn944 and Val945 in the dimer forced Cys947 out of the globular subdomain to contact Cys773 from the other molecule. In contrast, in the monomer, Cys773 and Cys947 extend in opposite directions and cannot make a disulfide bond (Fig. S1*E*).

Comparison of Human C-Ala with *A. fulgidus* C-Ala. As noted above, dimeric *A. fulgidus* C-Ala contains a long helical domain and a separate globular domain (Figs. 1*B* and 2*D*). The *A. fulgidus* C-Ala dimer is formed through a helix-loop-helix zipper (HLHZ) between the helical domains of the two partners (17). The structure of this dimer is unchanged in the context of the full-length dimeric *A. fulgidus* AlaRS (17, 19) (Fig. S2). However, after binding to one tRNA molecule, the globular domain of the full-length *A. fulgidus* AlaRS exhibits a conformational shift toward tRNA to contact the elbow region (Fig. S2). A comparison of human and archaeal C-Ala dimers shows that whereas archaeal C-Ala has a parallel dimer organization relying on HLHZ interactions, the *Hs* C-Ala has a

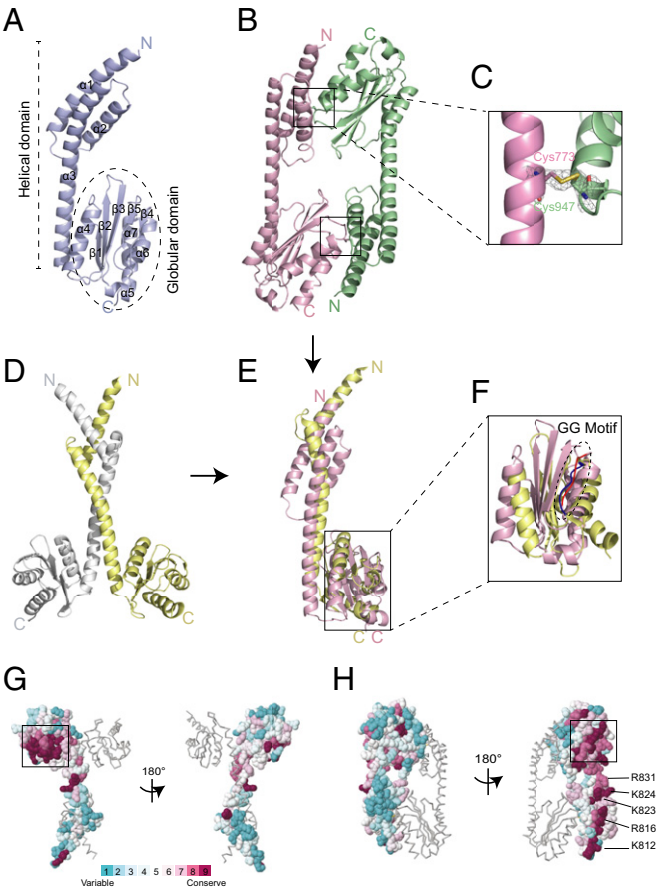


Fig. 2. Crystal structure of human C-Ala. (*A*) The crystal structure of monomeric human C-Ala. (*B*) The crystal structure of dimeric human C-Ala. Two disulfide bonds are shown in the black boxes. One molecule is shown in light purple, and the other is in pale green. (*C*) A 2Fo-Fc electron density map contoured at 1.5 σ. A disulfide bond was formed between Cys773 of one molecule and Cys947 of the other. (*D*) A dimeric form of *A. fulgidus* C-Ala. One molecule is shown in light yellow, and the other is in gray. (*E*) Superimposition of the monomers of human and *A. fulgidus* C-Ala. (*F*) A zoom-in view showing the superimposition of the globular domains of human and *A. fulgidus* C-Ala. Human C-Ala GG motif is colored in red, and *A. fulgidus* C-Ala GG motif is shown in blue. (*G*) Structure of *A. fulgidus* C-Ala with surface residues colored in accordance with evolutionary conservation (high, magenta; low, cyan) among amino acid sequences from different 150 C-Ala sequences. The boxed area shows the highly conserved GG motif. These figures were prepared using ConSurf server. (*H*) Structure of human C-Ala with surface residues colored in accordance with evolutionary conservation among amino acid sequences from different 150 C-Ala sequences. The positively charged residues (lysine or arginine), which is highly conserved from the helical domain, is labeled.

"head-to-tail" or antiparallel organization, with the globular domain of one monomer interacting with the helical domain of the other monomer and vice versa (Fig. 2 *B* and *D*).

Comparison of the monomers shows that whereas *A. fulgidus* C-Ala has two α-helices in the helical domain, C-Ala has an additional α-helix at the C-terminal end of the helical domain (Fig. 2*E*). Although these monomers' overall structures are similar, they did not superimpose well. When superimposing only the globular domains, the rmsd of Cα positions between human C-Ala and the C-Ala portion of *A. fulgidus* AlaRS is ∼5.8 Å, which means that the two domains are not well conserved as structures; however, we found that the glycine-rich or "GG" motifs (the GKGGG segment in human C-Ala and GSGGG segment in *A. fulgidus* C-Ala) are highly conserved (Fig. 2*F*). A comparison of the human and *A. fulgidus* C-Ala structures clearly

shows that a significant sequence divergence between the two species results in completely distinct architectures. Thus, our two crystal structures, which are the first of the eukaryotic AlaRS C-Ala domain, reveal a completely different dimerization interface that results in an antiparallel organization.

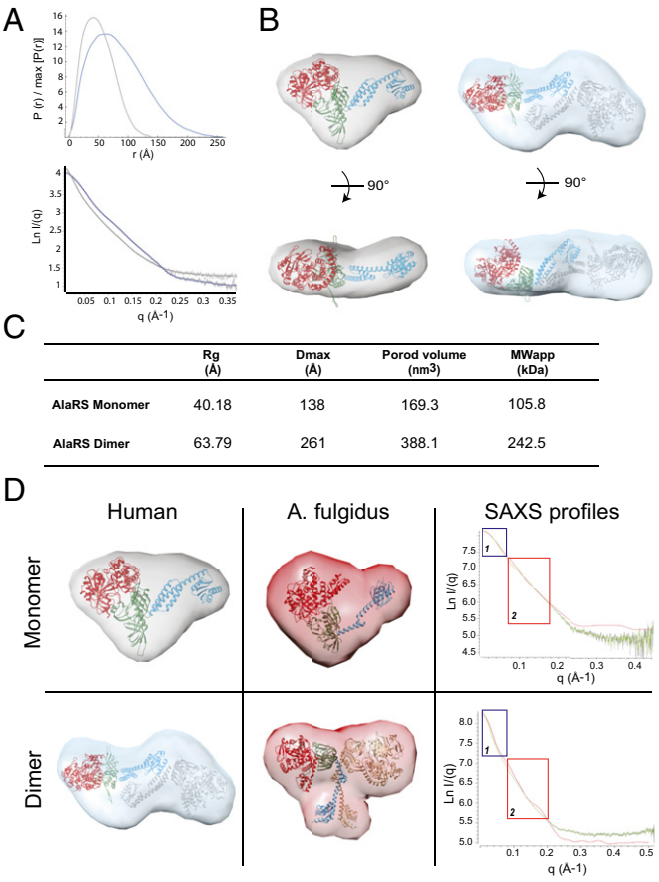
To investigate the evolutionary conservation of amino acids within C-Ala, we used ConSurf (23, 24) to determine the importance of each amino acid within the protein. We used both human and *A. fulgidus* structures as templates and compared them with 150 sequences of all homologous organisms displaying a sequence identity between 30% and 90%. For both structures, the region with the highest sequence conservation score is the GG motif that we previously identified in the globular domain (Fig. 2 *G* and *H*). Similarly, for both human and *A. fulgidus* templates, we found that all residues present at the dimer interface are highly variable and have the lowest conservation scores (1 and 2 on the scale) (Fig. 2 *G* and *H*). This low conservation explains how the HLHZ dimerization mode was lost during evolution, allowing the *Hs* C-Ala to form a structurally distinct dimer that may be associated with novel functions. Interestingly, within the eukaryote-specific α3 helix, a highly conserved motif occurs that contains structurally contiguous positively charged residues (K812, R816, K823, K824, R831, K834, and K846) (Fig. 2*H* and Fig. S1*F*). It is tempting to speculate that this highly conserved motif is involved specific functions in higher organisms.

Comparison of Human AlaRS with *A. fulgidus* AlaRS. To confirm that this distinct dimerization extends to the full-length AlaRS, we performed small-angle X-ray scattering (SAXS) on both the monomeric and dimeric forms of the full-length human AlaRS that was isolated from gel filtration (Fig. S3*A*). The shapes of the SAXS profiles and the corresponding Guinier plots obtained for both samples are independent of protein concentration, indicating the absence of significant aggregation (Fig. 3*A*). Guinier analysis in the low-*q* region gives radius of gyration (R_g) values of 40.2 Å for the monomeric form and 63.8 Å for the dimeric form (Fig. 3 *A* and *C*). The molecular masses calculated from the Porod volume are in agreement with what we expected for the monomeric form (105.8 kDa) and the dimeric form (242.5 kDa) of the protein (Fig. 3*C*).

We next used DAMMIF to perform ab initio shape reconstruction from the SAXS data. Several series of independent runs were carried out with no forced symmetry. All models were reproducible, with an average normalized spatial discrepancy (NSD) <1.0, indicating structurally similar solutions. Then the models resulting from 20 independent DAMMIF runs were superimposed and averaged using the DAMAVER suite to obtain a final averaged and filtered model (Fig. 3*B*).

We also generated a model of the full-length human AlaRS using the available human crystal structures for the aminoacylation and C-Ala domains, and also for the editing domain based on its structure in the *A. fulgidus* enzyme (25). The dimeric model was generated using the dimer interface of the crystal structure of the human C-Ala described herein. The resulting models for both the monomeric and dimeric human C-Ala fit well within the calculated SAXS envelopes (Fig. 3*B*). Importantly, a dimeric model based on the archaeal enzyme's dimerization interface cannot fit into the dimeric envelope, which clearly highlights the difference in the dimeric interface between human and archaeal orthologs.

We extracted the SAXS profile of *A. fulgidus* AlaRS from the previously solved crystal structure, and compared it with the experimental SAXS profile recorded for the *Hs* AlaRS for both monomer and dimer forms. The overall shapes of the two AlaRS monomers are quite similar (Fig. 3*D*), and superposition of the SAXS profile shows a good superimposition in the low- and medium-*q* value regions (Fig. 3*D*, boxes 1 and 2, respectively). The SAXS profile in the low-*q* region characterizes the global shape and size of the particle



	Rg (Å)	Dmax (Å)	Porod volume (nm³)	MWapp (kDa)
AlaRS Monomer	40.18	138	169.3	105.8
AlaRS Dimer	63.79	261	388.1	242.5

Fig. 3. Comparison of human AlaRS with *A. fulgidus* AlaRS. (*A*) The pairwise distance distribution function, *P*(*r*), of AlaRS-monomer (gray) and AlaRS-dimer (blue) (*Top*), and the theoretical scattering calculated from the average of 20 ab initio reconstructions (continuous lines, with AlaRS-monomer in gray and AlaRS-dimer in blue), plotted with the experimental scattering intensity curves (*Bottom*). The data are presented as the natural logarithm of the intensity. (*B*) The human full-length AlaRS model docked into the average ab initio SAXS envelope of the monomeric AlaRS (*Left*) and the dimeric full-length AlaRS model docked into the average ab initio SAXS envelope of the dimeric AlaRS (*Right*). The dimerization interface is based on the crystal structure of human dimer C-Ala. The aminoacylation domain is in red (PDB ID code 4XEM), the editing domain is in green, and C-Ala is in blue. (*C*) Summary of SAXS parameters. The R_g value was determined from the Guinier plot using AutoRg, and the maximum particle dimension (Dmax) and the Porod volume were calculated using GNOM. An estimate of the molecular weight was obtained by multiplying the Porod volume by 0.625. (*D*) Comparison of the human (*Left*) and *A. fulgidus* (*Middle*) envelopes for monomeric (*Top*) and dimeric (*Bottom*) full-length AlaRS. (*Right*) Alignment of the experimental SAXS profile for the human AlaRS (green) with the SAXS profile of the *A. fulgidus* AlaRS extracted from the crystal structure (red).

analyzed, whereas the medium-*q* region is related to local conformational differences. In the case of the dimeric proteins, the envelopes for the human AlaRS and the *A. fulgidus* AlaRS are clearly distinct, in agreement with a completely different dimerization interface between the AlaRSs of the two organisms. Moreover, superimposition of the SAXS profile clearly confirmed this difference, with a large deviation in both low- and medium-*q* value regions, thus supporting the different shapes of the proteins in solution.

Human C-Ala Binds DNA. Our observations led us to wonder why the human AlaRS evolved to form a new dimerization mode. Charging assays with both the monomer and dimer forms of human AlaRS revealed that this new dimerization mode does not affect tRNA binding or charging activity (Fig. S3*B*). Kinetic

analysis showed that k_{cat} was unchanged for human AlaRS vs. AlaRS-ΔC-Ala, whereas the K_m for tRNA was nominally improved (an ∼30% decrease) by the presence of the C-Ala domain (Table S2). In contrast, the effect of C-Ala is dramatic for the *E. coli* and *A. fulgidus* enzymes (16, 22). Therefore, it is likely that human C-Ala has little or no contact with bound tRNA, suggesting that it evolved to gain other functions.

From the crystal structure of the human C-Ala dimer, a disulfide bond is formed across the dimer interface between Cys773 and Cys947 (the numbering system of AlaRS). Whereas Cys947 was acquired in evolution by zebrafish and mouse and retained by human C-Ala, only *Hs* C-Ala added Cys773. Thus, under oxidative conditions, the C-Ala dimer should be further stabilized and become the dominant form. To obtain clues about the function of the *Hs* C-Ala dimer, we analyzed the surface charge of our crystal structure. Interestingly, one face of the C-Ala dimer had major positive patches, whereas the other face was rather neutral (Fig. 4*A*). Residues R793, K820, K823, K824, R831, K834, K839, and K846 from the helical domain combined with K869, K876, K898, K930, and K943 from the globular domain to form a positively charged groove that closely recapitulates a DNA-binding groove (Fig. 4*A*). Moreover, our conservation analysis with ConSurf showed that most of these residues (K823, K824, R831, K834, K846, K869, K876, K898, and K930) had a high conservation score (8 or 9) in the eukaryotic-specific helical motif (Fig. 2*H* and Fig. S1*F*).

To predict how DNA would bind to *Hs* C-Ala, we used the Discovery Studio ZDOCK program (26), a rigid-body protein-docking algorithm that uses a fast Fourier transform algorithm to speed search in translational space, to create a DNA-binding model (Fig. 4*B*). In this model, DNA binds to the human C-Ala dimer through the positive charges on surface of the dimer interface, with an unobstructed fit into the DNA-binding groove.

Although the *A. fulgidus* C-Ala dimer showed some positive patches on the surface, it did not exhibit any region that would likely form a DNA-binding groove (Fig. 4*C*). Given the knowledge that *E. coli* AlaRS can bind to and regulate expression of its own gene (27), we tested the DNA-binding activities of both *Hs* C-Ala and *Ec* C-Ala (Fig. 4*D*) through a DNA cellulose-binding assay. We incubated 2 μg of *Hs* C-Ala and 80 μg of *Ec* C-Ala with DNA cellulose and evaluated whether the proteins could be pulled down with the DNA cellulose. The proteins that bound to DNA cellulose were eluted in SDS sample buffer and run on an SDS/PAGE gel. The amount of protein bound to DNA was determined by silver staining. Although higher concentrations of *Ec* C-Ala were used, *Ec* C-Ala did not show any DNA binding, whereas *Hs* C-Ala showed robust DNA binding.

We also compared the DNA binding of *Hs* AlaRS (1 μg) and *Hs* C-Ala (1 μg), with *Hs* GlyRS (1 μg) used as a control (Fig. 4*E*). We determined the amount of protein bound to DNA through Western blot analysis with a His tag antibody. Whereas *Hs* AlaRS showed no binding to either cellulose (control) or

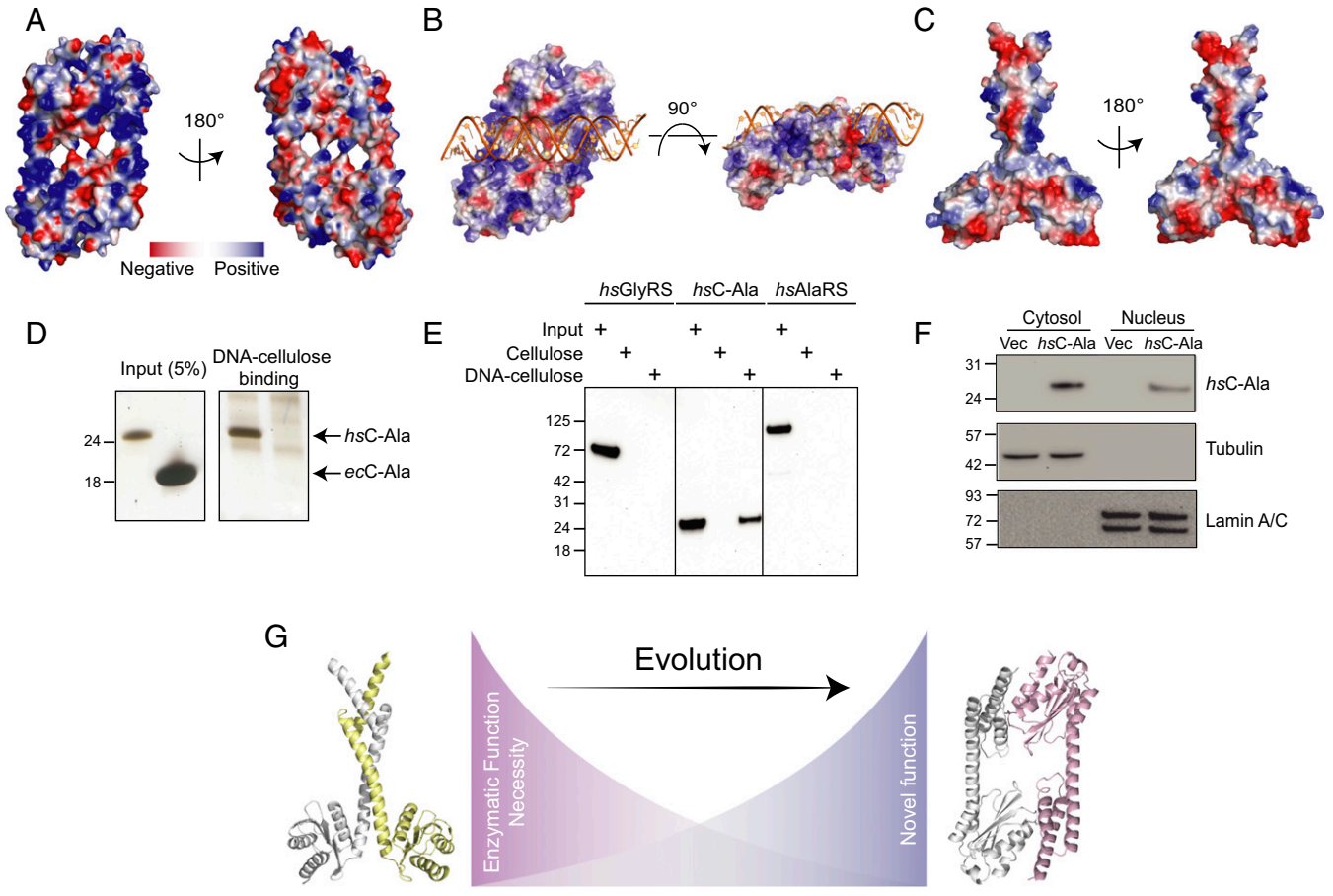


Fig. 4. Human C-Ala binds DNA. (*A*) Electrostatic surface views of human C-Ala dimer structure. (*B*) A DNA-binding model created by ZDOCK program showing that DNA binds to the human C-Ala dimer through the positive charges on surface of the dimer interface. (*C*) Electrostatic surface views of *A. fulgidus* C-Ala dimer structure. (*D*) DNA cellulose-binding assay of *Hs* C-Ala and *Ec* C-Ala. (*E*) DNA cellulose-binding assay of *Hs* GlyRS, C-Ala and AlaRS. (*F*) The nuclear distribution of *Hs* C-Ala. Lamin A/C and tubulin were used as nuclear (N) and cytoplasmic (C) markers, respectively. (*G*) The distinct dimerization mode for C-Ala during evolution.

DNA cellulose, *Hs* C-Ala showed strong binding to DNA cellulose, but not to cellulose alone. This result demonstrates that *Hs* C-Ala may have to be separated from full-length protein to have strong DNA-binding activities. The sequence-specificity of this DNA binding activity is under investigation. Because of its small size (23 kDa), we expected *Hs* C-Ala to enter the nucleus. This expectation was confirmed (Fig. 4*F*).

When we initiated the present study, AlaRS appeared to be an isolated example in which an appended domain was not introduced during eukaryote evolution to change the overall synthetase architecture and thereby introduce a new function. Here we show that AlaRS followed a different evolutionary path, by exploiting what appears to be a highly adaptable structure of a preexisting appended domain (C-Ala). This evolution resulted in a distinct dimerization mode for C-Ala, which in turn also changed the quaternary shape of *Hs* AlaRS (Fig. 4*G*). We raise the possibility that the change in architecture of both C-Ala and AlaRS resulted from the same evolutionary pressures that introduced one or more novel functions, as occurred by fusing new domains to the other 19 tRNA synthetases.

Methods

The crystal diffraction data were collected at 100 K on beamline 11-1 of the Stanford Synchrotron Radiation Lightsource (SSRL). All SAXS measurements were carried out at the SSRL on beamline BL4-2 at a working energy of 12.5 KeV. Detailed protocols for these experiments and additional procedures are described in *SI Methods*.

ACKNOWLEDGMENTS. We thank the staff at beamline 11-1 and BL4-2 of Stanford Synchrotron Radiation Lightsource (SSRL) for assistance in the data collection. We thank Dr. Shigeyuki Yokoyama for providing the *A. fulgidus* AlaRS gene plasmid. This work was supported by the National Cancer Institute (Grant CA92577), the National Foundation for Cancer Research, and US National Institutes of Health Grant R01 NS085092.

1. Guo M, Yang XL, Schimmel P (2010) New functions of aminoacyl-tRNA synthetases beyond translation. *Nat Rev Mol Cell Biol* 11(9):668–674.
2. Kim S, You S, Hwang D (2011) Aminoacyl-tRNA synthetases and tumorigenesis: More than housekeeping. *Nat Rev Cancer* 11(10):708–718.
3. Park SG, Schimmel P, Kim S (2008) Aminoacyl tRNA synthetases and their connections to disease. *Proc Natl Acad Sci USA* 105(32):11043–11049.
4. He W, et al. (2015) CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. *Nature* 526(7575):710–714.
5. Sampath P, et al. (2004) Noncanonical function of glutamyl-prolyl-tRNA synthetase: Gene-specific silencing of translation. *Cell* 119(2):195–208.
6. Han JM, et al. (2012) Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway. *Cell* 149(2):410–424.
7. Bonfils G, et al. (2012) Leucyl-tRNA synthetase controls TORC1 via the EGO complex. *Mol Cell* 46(1):105–110.
8. Yao P, Fox PL (2013) Aminoacyl-tRNA synthetases in medicine and disease. *EMBO Mol Med* 5(3):332–343.
9. Lo WS, et al. (2014) Human tRNA synthetase catalytic nulls with diverse functions. *Science* 345(6194):328–332.
10. Guo M, Schimmel P (2013) Essential nontranslational functions of tRNA synthetases. *Nat Chem Biol* 9(3):145–153.
11. Jia J, Arif A, Ray PS, Fox PL (2008) WHEP domains direct noncanonical function of glutamyl-Prolyl tRNA synthetase in translational control of gene expression. *Mol Cell* 29(6):679–690.
12. Rho SB, et al. (1999) Genetic dissection of protein–protein interactions in multi-tRNA synthetase complex. *Proc Natl Acad Sci USA* 96(8):4488–4493.
13. Ko YG, Park H, Kim S (2002) Novel regulatory interactions and activities of mammalian tRNA synthetases. *Proteomics* 2(9):1304–1310.
14. Chang CY, Chien CI, Chang CP, Lin BC, Wang CC (2016) A WHEP domain regulates the dynamic structure and activity of *Caenorhabditis elegans* glycyl-tRNA synthetase. *J Biol Chem* 291(32):16567–16575.
15. Grice SJ, et al. (2015) Dominant, toxic gain-of-function mutations in gars lead to non-cell autonomous neuropathology. *Hum Mol Genet* 24(15):4397–4406.
16. Guo M, et al. (2009) The C-Ala domain brings together editing and aminoacylation functions on one tRNA. *Science* 325(5941):744–747.
17. Naganuma M, Sekine S, Fukunaga R, Yokoyama S (2009) Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proc Natl Acad Sci USA* 106(21):8489–8494.
18. Yachdav G, et al. (2014) PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* 42(Web Server issue): W337–343.
19. Naganuma M, et al. (2014) The selective tRNA aminoacylation mechanism based on a single G•U pair. *Nature* 510(7506):507–511.
20. Jasin M, Regan L, Schimmel P (1983) Modular arrangement of functional domains along the sequence of an aminoacyl tRNA synthetase. *Nature* 306(5942):441–447.
21. Jasin M, Regan L, Schimmel P (1984) Dispensable pieces of an aminoacyl tRNA synthetase which activate the catalytic site. *Cell* 36(4):1089–1095.
22. Fukunaga R, Yokoyama S (2007) Crystallization and preliminary X-ray crystallographic study of alanyl-tRNA synthetase from the archaeon *Archaeoglobus fulgidus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63(Pt 3):224–228.
23. Celniker G, et al. (2013) ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Isr J Chem* 53(3-4):199–206.
24. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38(Web Server issue):W529–533.
25. Sun L, et al. (2016) Evolutionary gain of alanine mischarging to non-cognate tRNAs with a G4:U69 base pair. *J Am Chem Soc* 138(39):12948–12955.
26. Pierce BG, et al. (2014) ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 30(12):1771–1773.
27. Putney SD, Schimmel P (1981) An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription. *Nature* 291(5817):632–635.
28. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
29. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
30. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64(Pt 1):112–122.
31. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
32. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
33. Vagin A, Teplyakov A (1997) MOLREP: An automated program for molecular replacement. *J Appl Cryst* 30:1022–1025.
34. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* 36: 1277–1282.
35. Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst* 42(Pt 2):342–346.
36. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst* 36:860–864.
37. Beebe K, Waas W, Druzina Z, Guo M, Schimmel P (2007) A universal plate format for increased throughput of assays that monitor multiple aminoacyl transfer RNA synthetase activities. *Anal Biochem* 368(1):111–121.