Published in final edited form as:

J Phys Chem B. 2010 December 16; 114(49): 16273–16279. doi:10.1021/jp1082517.

Packaging HIV Virion Components Through Dynamic Equilibria of a Human tRNA Synthetase

Min Guo¹, Ryan Shapiro¹, Garrett M. Morris², Xiang-Lei Yang¹, and Paul Schimmel¹

¹ The Skaggs Institute for Chemical Biology and the Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

² Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Abstract

Aminoacyl tRNA synthetases, components of the translation apparatus, have alternative functions outside of translation. The structural and mechanistic basis of these alternative functions is of great interest. As an example, reverse transcription of the HIV genome is primed by a human lysine-specific tRNA (tRNA^{Lys3}) that is packaged (into the virion) by the HIV gag protein with lysyl-tRNA synthetase (LysRS). Not understood is the structural basis for simultaneous packaging of tRNA^{Lys3}, LysRS and Gag. Here *ab initio* computational methods, together with our recent high resolution 3-D structure of human LysRS, produced an energy-minimized model where Gag, tRNA^{Lys}, and LysRS form a ternary complex. Interestingly, the model requires normally homodimeric LysRS to dissociate into a monomer that bridges between Gag and tRNA^{Lys3}. Earlier experiments of others, and new experiments presented here, which tested an engineered dissociated form of LysRS, were consistent with the *ab initio* ‘bridging-monomer’ model. The results support an emerging theme that alternative functions of tRNA synthetases may come in part from protein surfaces exposed by dynamic equilibria.

Introduction

The algorithm of the genetic code is established in the first reaction of protein synthesis, where aminoacyl tRNA synthetases (AARSs) catalyze the aminoacylation reaction that fuses each amino acid to its cognate tRNA.¹ With its essential role, AARSs appeared early and were universally retained throughout evolution. The enzymes (one for each of the 20 amino acids) are divided evenly into two classes of 10 members for each. This classification is based on conserved structural features for the catalytic unit of each member of the same class. In addition, while sharing a common architecture for the catalytic core, enzymes in the same class have different quaternary structures. These include α (monomer), α_2 homodimer, and $\alpha_2\beta_2$ (heterotetramer). The rationale for these different quaternary forms is not known.

Here we focus on human LysRS, an α_2 homodimer that has multiple ex-aminoacylation functions². For example, LysRS interacts with microphthalmia-associated transcription factor (MITF)³ and upstream stimulatory factor (USF2),⁴ as part of a mechanism for transcriptional control of target genes.⁵ As another example, LysRS is packaged into the HIV virion *via* its interaction with the Gag protein.⁶ The HIV packaging of LysRS into the virion also includes a specific cognate tRNA isoacceptor--tRNA^{Lys3}.⁷ Because the packaged tRNA^{Lys3} is the primer for reverse transcription of the HIV genome, LysRS appears to have

an important role in the life cycle of the virus. Not understood is how this homodimeric tRNA synthetase can simultaneously interact with the Gag protein and with tRNA^{Lys3}.

Here, based on knowledge of the human LysRS structure determined earlier by us, and using basic thermodynamic principles,^{8,9} we calculated an *ab initio* LysRS-Gag docking model. We considered the enzyme in both dimer and monomer states. From this analysis, the LysRS-Gag interface was seen to overlap the covered surface of the α_2 LysRS dimer interface. This result raised the possibility that the dynamic equilibrium of LysRS between monomer and dimer states could direct the protein from aminoacylation (which requires the dimer) to HIV packaging (monomer, which is inactive for aminoacylation). We then designed a monomer form LysRS, with its potential Gag binding interface remaining intact. Our design switched the monomer-dimer equilibrium of LysRS toward the monomer state, and the capacity of this stabilized monomer for binding tRNA^{Lys3} was then investigated.

Methods

Constructs and Protein Preparation

All constructs were generated using a standard polymerase chain reaction-based cloning strategy. Full length LysRS (1–597) and its truncated form (70–584) were expressed in the bacterial strain BL21 (DE3) CodonPlus using the pET20b vector (Novagen). Both proteins contained a 6 × His tag at their C-terminus, and were purified to homogeneity by a Ni-NTA affinity column (Qiagen, Valencia, CA) and a Q high performance column (GE Healthcare, Little Chalfont, Buckinghamshire, U.K.). All proteins were concentrated in 5 mM Tris-HCl buffer, pH 8.0, 50 mM NaCl and 5 mM beta-mercaptoethanol.

Engineering of LysRS

Mutations on the truncated form of human LysRS (70–584) were introduced by the QuikChange Mutagenesis PCR method (Stratagene) and verified by DNA sequencing. The *E. coli* *flaB* gene encoding Flavodoxin -2 (UniprotKB: P0ABY4) was amplified from *E. coli* K-12 genome DNA. It was then inserted into either empty pET20b vector, or into the human LysRS expression plasmid using the QuikChange method modified for large insertions. All proteins were expressed and purified identically to native human LysRS.

Gel Filtration Chromatography

For each assay, 500 μ l of purified truncated human LysRS (70–584) and its mutants were applied to a Superdex 200 chromatography column (GE Healthcare, 10/300 GL) in a buffer containing 25 mM HEPES, pH7.5, 150 mM NaCl and 5 mM beta-mercaptoethanol. The peak fractions were analyzed by SDS-PAGE and visualized by Coomassie blue staining.

Crosslinking Assay

A T285C mutation was introduced into native human LysRS (70–584) and *Flavo*-LysRS (70–584). All proteins were expressed and purified as mentioned above, except that no beta-mercaptoethanol was added. The purified proteins were diluted to 1mg/ml in 25 mM HEPES (pH7.5, 150 mM NaCl) and incubated for 15 minutes at 4 °C with or without 10 mM DTT (dithiothreitol). Proteins samples (10 μ l) were mixed with an equal amount of 2× loading buffer (without reducing reagent) and boiled for 10 min before loading onto a SDS-PAGE gel.

Pull-down Assay

HIV-1 Gag Δ p6 template was a kind gift from Dr. A. Rein. The region 1–448 of Gag Δ p6 (encodes matrix, capsid and nucleocapsid) was subcloned into a homemade GST fusion

vector, pHisGSTtev. An octapeptide StrepII-tag (sequence WSHPQFEK) was inserted at the N-terminus of GagAp6. The HisGSTtev-StrepII-Gag p6 fusion protein was expressed and purified identically to the procedure used for human LysRS. The N-terminal HisGST was cleaved by TEV protease (1:100) at 4°C overnight and removed by passing through a Ni-NTA affinity column. The purified StrepII-GagAp6 protein was bound onto 50µl Strep-Tactin beads (IBA, Germany) and washed. Then 200µl of 20µM WT LysRS, Flavo-LysRS or *E. coli* Flavodoxin was added to the beads and incubated at 4°C for 2 hours. Unbound proteins were washed with standard washing buffer (100mM Tris-HCl (pH8.0), 150 mM NaCl, 1 mM EDTA). Bound proteins were then eluted with elution buffer (2.5 mM desthiobiotin in washing buffer). The samples were checked by Western blot using anti-His antibody.

EMSA Assay

Human tRNA^{Lys3} was prepared by in vitro transcription based on previously described methods.¹⁰ The annealed human tRNA^{Lys3} was then radiolabeled with ³²P at the 3'-end as described previously.¹¹ For each complex formation reaction, about 10,000 cpm of tRNA (diluted with cold tRNA^{Lys3} to a final concentration of 0.1 µM) was incubated for 10 min on ice with variable concentrations of human LysRS or Flavo-LysRS (full-length, 1–597) in a total volume of 20µl containing 20 mM HEPES (pH 7.5), 20 mM KCl, 5 mM MgCl₂, and 2 mM DTT. Glycerol was added to a final concentration of 5% in each sample prior to loading on a native 6% 29:1 polyacrylamide gel. The gel ran at 100 V for 1.5 hr at 4 °C in 0.5× TBE, was fixed by 7% HOAC for 5 min at 4 °C, and then dried onto analytical filter paper (Schleicher & Schuell).

Docking

Coordinates of Gag-CA-CTD were taken from the crystal structure of CTD (*pdb ID*: 1A80). Initial docking of CTD on both monomer and dimer structures of human LysRS (*pdb3BJU*) was estimated using ZDOCK. The solutions were clustered by ClusPro with binding energies calculated by both FastContact and Dcomplex. Then a blind-docking of CTD on the monomer structure of human LysRS was performed with the default parameter on 3 docking programs, HEX, ZDOCK and Dot.^{12–14} The binding site was screened by docking the CA-CTD domain on all surfaces of the human LysRS structure (*pdb3BJU*, chain A). Next, we refined the docking of the CA-CTD domain in a 50-Å size cubic space centering on the dimer interface. A thorough rigid-body docking was carried out using AutoDock.¹⁵ After searching with the maximum grids and clustering, the large cluster with the lowest energy showed apparent separation from all the others. Then, this lowest energy solution was subjected to molecular dynamic refinement in four steps by CNS¹⁶: rigid-body torsion angle dynamics of two groups (LysRS and CA-CTD, 500 MD cooling steps from 2000 K with a 8 fs time step), rigid-body torsion angle dynamics of three groups (LysRS_N, LysRS_C and -CA, 500 MD steps from 2000K), semi-flexible simulated annealing (1000 MD steps from 1000 K with 4 fs time steps) with side chains of the interface residues, and a final semi-flexible simulated annealing (1000 MD steps from 1000K to 50K with 2fs time steps), where both side chains and backbones of the interface residues were allowed to move to allow for conformational arrangement. After the MD refinement, the interaction energy (sum of E_{elec} , E_{vdw} , E_{ACS}) decreased from -10691.001 to -11788.829 kcal/mol.

Results

Ab initio docking of Gag CA-CTD to the human LysRS monomer

The structure of human LysRS has a class II C-terminal catalytic domain, preceded by an anticodon-binding domain. Support has been obtained for human LysRS binding to the C-terminal domain (CTD) of the capsid region (CA) of the HIV-1 Gag protein.^{10,17} A model

of the human LysRS/CA-CTD complex has been presented by others.¹⁸ This model proposed a close interaction of the CA-CTD helix 4 (H4) with one face of helix H7 of human LysRS. In this instance, the structure of *E. coli* LysRS was used to generate a model for the human counterpart. A rigid-body docking was then carried out using CA-CTD and the modeled human LysRS structure. Importantly, the final model was selected for its consistency with the experimental mapping data.

In our work, we approached the problem in the inverse way. At the start of the analysis, we noted that, when comparing our recently solved crystallographic structure of human LysRS with that of *E. coli* LysRS, the orientation of the N-terminal anticodon domain, and the locations and kinds of surface residues, were substantially different.¹⁹ Specifically, the N-terminal anticodon-binding domain rotates about 5° between these two structures, and only 25% of the residues are conserved between the two surfaces. These observations gave strong motivation to start with an *ab initio* analysis of how Gag-CA-CTD bound with human lysyl-tRNA synthetase. We could then test this model with published and additional experimental results (see below).

Using Zdock, we first performed systematic rigid-body docking of Gag-CA-CTD with both the monomer and, separately, the dimeric structure of human LysRS. Interestingly, in the top 10 clusters of CTD dockings to dimeric LysRS, nine showed CTD binding to only one subunit of dimeric LysRS. (For one cluster, CTD bound to both subunits of dimeric LysRS, but with only about half the binding energy compared to the other top hits.) Thus, binding of CTD of the two subunits of dimeric LysRS is unfavorable. When CTD was docked on the monomeric structure of human LysRS, the docking sites were localized predominantly to the exposed residues used to form the dimer interface. This interface is covered and, therefore, inaccessible in dimeric LysRS. Thus, the congruence of the primary docking trial suggested CTD has a tendency to bind to monomeric human LysRS.

To confirm this result, a few well-developed computational docking programs were tested to achieve an unbiased fit of monomeric LysRS with Gag-CA-CTD. Complete searches of all orientations between the two proteins were done by systematically rotating and translating a moving Gag-CA-CTD with a stationary LysRS. Interaction energies were computed as the sum of van der Waals and electrostatic terms (see Methods). Significantly, docking Gag-CA-CTD to monomeric human LysRS by each of 3 programs (HEX, Zdock, Dot) identified one common Gag-CA-CTD binding site. This common site fell within 90–95% of the solutions represented in the 200 top-ranked solutions for all 3 programs (Fig. 1A). For this common docking site, Gag-CA-CTD docking solutions were located in a 35-Å wide cavity of LysRS formed by the N-terminal anticodon-binding domain and the C-terminal aminoacylation domain. This cavity is a major part of the surface needed to form the LysRS homodimer, and constitutes ~60% of the entire dimeric interface between individual LysRS monomers. The bottom of this cavity forms the largest hydrophobic patch of the LysRS surface. The extensive overlap of these docking solutions from independent docking algorithms strongly suggests that this cavity is the preferred site for binding of LysRS to Gag-CA-CTD.

To enumerate all possible binding orientations, we utilized Autodock, which is often used for accurate placement of small molecules on proteins.¹⁵ This second step ‘Autodock analysis’ was performed in a defined cubic space (50 Å × 50 Å × 50 Å) centered on the inter-domain cavity (formed between aminoacylation and anticodon-binding domains) that was seen in the profile of the initial general searches (Fig. 1B). Significantly, Autodock gave a cluster of docking orientations having the lowest energy, and well separated from other clusters. The largest grouping is the 2nd lowest energy cluster, located within a broad but shallow energy state. In contrast, the lowest energy cluster sits in a deep but narrow energy

funnel. Thus, a specific orientation of the two molecules defines the lowest energy interaction.

Molecular dynamic simulations suggest a conformational change needed for LysRS/Gag-CA-CTD complex formation

The lowest energy solution of Autodock was subjected to further molecular dynamics refinement. The docking solution of lowest energy positioned Helix3 and Helix4 of Gag-CA-CTD into contact with the hydrophobic patch (231–245) in the aforementioned groove of monomeric human LysRS (Fig. 1C). The Gag-CA-CTD was located close to the anticodon-binding domain of LysRS. Inspection of this model revealed that contacts between Gag-CA-CTD and motif 1 helix (238–260) in LysRS would be possible, provided that the conformation of the anticodon-binding domain is flexible. Because flexibility was suggested by the orientation of the anticodon domain of human LysRS being rotated by 5° compared to that of its *E. coli* ortholog, we continued to optimize this model. The first rigid-body refinement of the complex was achieved by releasing constraints on the entire N-terminal anticodon-binding domain (of LysRS) and on Gag-CA-CTD. A molecular dynamics (MD) experiment was then carried out by gradually relaxing the interface residues and the conformation of the anticodon-binding domain. The model converged during the MD refinement to a state that improved the favorable interaction energy by −1.1 kcal/mol compared to the initial model. Remarkably, although the anticodon domain only shifted 0.9 Å, the Gag-CA-CTD domain shifted ~3 Å towards a helix that is present in the aminoacylation domain of all class II tRNA synthetases (motif-1 helix) (Fig. 1C). The newly formed contacts between Gag-CA-CTD and the aminoacylation domain compensated for the loss of partial contacts between the Gag-CA-CTD and the anticodon-binding domain (Fig. 1D, 1E). This result suggested that a conformational change induced by Gag-CA-CTD binding is energetically favored.

Model of human LysRS/tRNA^{Lys3}/Gag-CA ternary complex

With the LysRS/CA-CTD docking model as a reference, we constructed a model of the tRNA^{Lys3}/LysRS/Gag-CA-CTD complex. For this, we used structural information on *E. coli* AspRS in complex with tRNA^{Asp}. (LysRS is more closely related to AspRS than to any other tRNA synthetase.) The tRNA from the *E. coli* AspRS-tRNA^{Asp} complex structure (*pdb1ASY*) was built onto human LysRS by simple superposition of the respective catalytic domains (Fig. 2). To complete the full structure of human LysRS, we added the N-terminal eukaryotic extension that is thought to bind nonspecifically to the acceptor-T ψ C stem-loop domain of tRNA.^{20,21} The N-terminal portion of the extension is predicted as a long helix that is positioned on the elbow region of the modeled tRNA. For this positioning, we used the functionally similar A1 domain of *Thermus thermophilus* PheRS bound to tRNA^{Phe} (*pdb1E1Y*). This strongly dipolar lysine-rich helix has a positive electrostatic potential extending out from one side of the protein to form part of the RNA-binding site. We then manually docked the N-terminal domain (NTD) of the capsid protein onto the complex by superimposing Gag-CA-CTD from the known structure of the Fab-capsid complex (*pdb1E6J*).²² Although the linker between N- and C- domains of the capsid protein are flexible, and the capsid protein can adopt multiple conformations that are important for virion assembly,²³ Gag-CA-CTD was positioned far from the LysRS-tRNA^{Lys3} complex-- a position with least impact on the binding interface (Fig. 2). Based on this model, it is easy to understand why Gag-CA-CTD alone is sufficient to bind to LysRS, and why the rest of Gag is dispensable for binding to LysRS.¹⁷

Conversion of dimeric human LysRS to monomer

Because the LysRS/Gag-CA-CTD interface overlaps with the core of the LysRS homodimer interface, a monomer form of LysRS is required to access that interface. Given that dimeric

LysRS and the modeled LysRS/Gag-CA-CTD complex share the interface, we aimed at designing a disruption of the LysRS dimer interface without affecting the interface of LysRS with Gag-CA-CTD. Based on the model, we started with a series of mutations on the surface of LysRS that form part of the dimer interface but are outside of the part of the interface that binds to Gag-CA-CTD (Fig. 3A, 3B). Gel filtration analysis showed that none of these mutant proteins behaved as monomers and some formed aggregates (see below). Possibly this lack of monomer-disruption occurs because the dimer interface of LysRS is so extensive (5800 Å², 24% of the monomer surface). This extensive interface is reflected in the low K_d for dissociation of the dimer (~2 nM).¹⁹

Because direct mutations on the surface of LysRS failed to produce a monomer, we applied another approach. We imagined that a large globular insertion (at least 30 Å in diameter) on the periphery of the dimer interface should provide enough steric hindrance to block dimer formation. We then searched the known structures in the protein database with the following criteria: 1. Not a membrane protein (so as to not introduce an extra hydrophobic surface for protein-protein interactions); 2. A globular 3D structure; 3. 30–40 Å dimension; 4. N-terminus and C-terminus located in close proximity (to minimize a structural change when inserted it into a single site). The structure of *E. coli* Flavodoxin (*pdb1ahn*) fits all criteria. The gene encoding this 180 amino acid protein was then amplified, cloned and inserted at the G310 position of human LysRS (Fig. 3C). The resulting recombinant *Flavo*-LysRS fusion protein was expressed and purified. Gel filtration, at a loading concentration of 2 μM, showed that *Flavo*-LysRS ran mostly as a monomer (Fig. 3D).

FMN (flavin mononucleotide) bound to native *E. coli* flavodoxin has a characteristic absorbance centered at 466 nm, thus giving a bright yellow color.²⁴ Purified *Flavo*-LysRS had a similar absorbance spectrum, consistent with the flavodoxin domain of the fusion protein being in the native conformation (Fig. 4A). To verify the monomeric state of *Flavo*-LysRS seen by gel filtration, we carried out a crosslinking assay in solution. For this purpose, we noted that two b-hairpins in the aminoacylation domain of each monomer form an anti-parallel hydrogen bonded b-sheet interaction, as part of the dimer interface. The Thr285 side chains of these hairpins on the two subunits are 4.5 Å apart (Fig. 4B). Because this spacing could accommodate a disulfide crosslink, a T285C mutation was introduced into both native- and *Flavo*-LysRS. With this T285C substitution, native LysRS-T285C spontaneously formed a covalent dimer during purification; in contrast, native LysRS did not. Addition of 5 mM dithiothreitol (DTT) converted the covalent dimer into a monomer on the SDS-PAGE gel (Fig. 4B). In contrast, for WT and T285C *Flavo*-LysRS, both forms showed the same pattern, with little crosslinked dimer found on the SDS-PAGE gel (Fig. 4B). Thus, the flavodoxin insertion efficiently disrupted the tight dimerization of LysRS (Fig. 3C,D).

Monomeric *Flavo*-LysRS retains Gag and tRNA^{Lys3} binding

Based on our LysRS/tRNA^{Lys3}/Gag-CA model, the flavodoxin was designed to insert at a site that would give minimal steric interference to complex formation. In order to check the feasibility of the model, we then experimentally examined the binding of *Flavo*-LysRS to Gag (Fig. 5A). We used a GagΔp6 construct fused with the StrepII -tag and bound it to Strep-Tactin beads as bait. From the pull-down assay, monomeric *Flavo*-LysRS bound to Gag, while flavodoxin alone showed no binding to Gag. Although in this particular assay *Flavo*-LysRS seemed to bind better to Gag than did LysRS, more quantitative studies need to be carried out in the future.

Because packaging *in vivo* of tRNA^{Lys3} into the HIV virion requires LysRS^{3,7}, we carried out an electrophoretic mobility shift assay (EMSA) to investigate binding of tRNA^{Lys3} to LysRS. The gene for tRNA^{Lys3} was transcribed *in vitro* and labeled with ³²P at the 3'-end.

Using the EMSA method, both native- and *Flavo*-LysRS bound tRNA^{Lys3} with similar affinities (Fig. 5B, 5C). (*Flavo*-LysRS had a slightly weaker affinity of 1.21 μ M than did native LysRS (0.87 μ M, Fig. 5D).) Therefore, monomeric *Flavo*-LysRS retains the ability to bind tRNA^{Lys3}.

Discussion

As stated in the Introduction, our goal was to establish an *ab initio* model for the LysRS/Gag-CA-CTD/tRNA^{Lys} interaction, which could then be examined in the light of published experimental data and of new experiments designed by us. Although docking subunits that undergo conformational changes (as seen here) is challenging, the prediction of protein-protein docking can now create structures with atomic-level accuracy.²⁵ Here, with only structural information, we used a three-stage approach to predict the binding of human LysRS with Gag-CA-CTD (Fig. 1). After docking by several independent programs, we were able to converge to a defined space. We then locally sampled the six-dimensional space of LysRS/CA-CTD structures for the strongest binding (Autodock). Finally, we subjected the lowest free energy solution to a stepwise MD refinement, focusing on the flexibility of each domain and on all interface residues. Again based on the predominance of both energy and cluster populations, each of the other three programs (HEX, Zdock, Dot) yielded binding sites around the dimer interface of LysRS. Thus, the inter-domain cavity, which is exposed in monomeric LysRS, forms a natural binding site for Gag-CA-CTD.

The refined docking model predicted that H3, H4 and the C-loop of CA-CTD form the binding interface. Consistent with this prediction, *in vitro* immunoprecipitation experiments showed that removal of H4 (205–231) abolished the ability of CA-CTD to bind to LysRS.¹⁷ In addition, immunoprecipitation assays with whole cell lysates showed that, without the C-terminal half of CA-CTD (191–231) that spans the C-loop, the partially deleted Gag loses its interaction with endogenous LysRS.¹⁷ Further dissection showed a peptide constituting H4 and the C-terminal loop (211–227) in CA can bind to LysRS with an affinity similar to that of CA-CTD.¹⁸ In our model, the C-terminal half of the CA-CTD domain, consisting of helices H3 (196–205) and H4 (211–218), was positioned into the deeper part of the LysRS inter-domain cavity (Fig. 1E). Importantly, when the C-terminal loop (220–231) of CA-CTD was added into the docking model, we saw potential contacts between this loop and LysRS (Fig. 2). These predicted contacts might explain why the loop residues gave the strongest shifts in an NMR titration of LysRSs with CA-CTD.¹⁸

The docking model also showed that the minimal CA-CTD binding site is located within a segment at the N-terminal side of motif 1 (228–255) of LysRS (Fig. 1C–E). *In vitro* assays of others showed that residues 1–308 and 220–597 of LysRS each bind to Gag-CA-CTD with an affinity similar to that of full-length LysRS.¹⁷ Further, *in vivo* packaging assays also showed that residues 1–259 and 208–597 can be efficiently incorporated into Gag-containing viral-like-particles.¹⁷ Collectively, the experimental assays mapped out the minimal binding region of LysRS to residues 220–259. This minimal binding region fits well with our predicted binding site of 228–255 (Fig. 1E).

Based on our docking model, the full-length LysRS/tRNA/Gag-CA complex has the C-terminal end of the Gag protein located on the same side as bound tRNA on LysRS. In contrast, the N-terminal end of Gag is on the opposite side. This orientation would make the packaged tRNA easy to be accessed by the C-terminal nucleocapsid part of Gag, for productive packaging of tRNA^{Lys3} and annealing to the HIV genome.^{26,27}

Our docking results also showed that H3, H4 of CA-CTD bind to the inter-domain cavity of LysRS monomer (Fig. 1E). It has been suggested that the available interaction sites on CA-

CTD are limited, with only one site being suitable as a target for effective inhibitor binding.²⁸ A 12mer peptide was reported to bind to a groove spanning H1, H2 and the top of H3 and H4. However, our model indicates that the bottom side of H3 and H4 could form an interaction interface with human LysRS and could be used as a target for a drug. Because H3, H4 and the C-terminal loop of CA-CTD are critical for HIV virion assembly,²⁹ an inhibitor covering the 'bottom side' of H3 and H4 may not only inhibit LysRS binding and tRNA^{Lys} incorporation, but would also be effective for blocking virion assembly.

Although monomeric LysRS, like other dimeric class II tRNA synthetases that have been tested, is inactive for aminoacylation¹⁰, our results showed that disruption of LysRS dimerization seems not to have a major effect on tRNA binding (Fig. 4E). The N-terminal helix motif specific to eukaryotic LysRS plays a critical role for capturing tRNA^{Lys} and contributes significantly to the binding.^{20,30} On the other hand, based on the crystal structure of the bacterial LysRS-tRNA^{Lys} complex, tRNA^{Lys} is mainly anchored through the interaction of the anticodon stem-loop with the anticodon-binding domain, with minimal interaction being seen with the catalytic domain of the synthetase.³¹ Because tRNA binding affinity of monomeric Flavo-LysRS is close to that of the native dimeric LysRS, monomeric LysRS can efficiently capture tRNA. (Other interactions between tRNA^{Lys3} with Gag-Pol may further stabilize the ternary complex.) Thus, the monomeric Flavo-LysRS described here could be a valuable tool for studying the mechanism of tRNA^{Lys3} incorporation into the HIV virion.

As stated earlier, many class II tRNA synthetases have dimeric quaternary structures. It is these quaternary forms that are catalytically active, while the monomeric forms are inactive. At the same time, many tRNA synthetases also have novel functions beyond aminoacylation. Because most class II AARSs share a similar dimer interface (formed by the motif-1 signature sequence),³² our results suggest the possibility that the monomer-dimer equilibrium is a mechanism to switch between aminoacylation (dimer) and their novel (monomer) functions.

Acknowledgments

This work was supported by grants GM 15539, GM 23562 and U54RR025204 from the National Institutes of Health, by grant CA92577 from the National Cancer Institute and by a fellowship from the National Foundation for Cancer Research. We thank Professor Karin Musier-Forsyth and Dr. Michael Ignatov for helpful comments on the manuscript.

References

1. Ibba M, Soll D. *Annu Rev Biochem.* 2000; 69:617. [PubMed: 10966471]
2. Nechushtan H, Kim S, Kay G, Razin E. *Adv Immunol.* 2009; 103:1. [PubMed: 19755181]
3. Lee YN, Nechushtan H, Figov N, Razin E. *Immunity.* 2004; 20:145. [PubMed: 14975237]
4. Lee YN, Razin E. *Mol Cell Biol.* 2005; 25:8904. [PubMed: 16199869]
5. Yannay-Cohen N, Carmi-Levy I, Kay G, Yang CM, Han JM, Kemeny DM, Kim S, Nechushtan H, Razin E. *Mol Cell.* 2009; 34:603. [PubMed: 19524539]
6. Kleiman L, Cen S. *Int J Biochem Cell Biol.* 2004; 36:1776. [PubMed: 15183344]
7. Cen S, Javanbakht H, Niu M, Kleiman L. *J Virol.* 2004; 78:1595. [PubMed: 14722314]
8. Alberty, RA. *Thermodynamics of biochemical reactions.* Wiley; Hoboken, NJ: 2003.
9. Silbey, RJ.; Alberty, RA.; Bawendi, MG. *Physical chemistry.* 4. Wiley; Hoboken, NJ: 2005.
10. Kovalski BJ, Kennedy R, Hong MK, Datta SA, Kleiman L, Rein A, Musier-Forsyth K. *J Biol Chem.* 2006; 281:19449. [PubMed: 16702215]
11. Guo M, Chong YE, Beebe K, Shapiro R, Yang XL, Schimmel P. *Science.* 2009; 325:744. [PubMed: 19661429]

12. Ritchie DW, Kozakov D, Vajda S. *Bioinformatics*. 2008; 24:1865. [PubMed: 18591193]
13. Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, Weng Z. *Proteins*. 2007; 69:511. [PubMed: 17623839]
14. Mandell JG, Roberts VA, Pique ME, Kotlovyy V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. *Protein Eng*. 2001; 14:105. [PubMed: 11297668]
15. Huey R, Morris GM, Olson AJ, Goodsell DS. *J Comput Chem*. 2007; 28:1145. [PubMed: 17274016]
16. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. *Acta Crystallogr D Biol Crystallogr*. 1998; 54:905. [PubMed: 9757107]
17. Javanbakht H, Halwani R, Cen S, Saadatmand J, Musier-Forsyth K, Gottlinger H, Kleiman L. *J Biol Chem*. 2003; 278:27644. [PubMed: 12756246]
18. Kovalski BJ, Kennedy R, Khorchid A, Kleiman L, Matsuo H, Musier-Forsyth K. *J Biol Chem*. 2007; 282:32274. [PubMed: 17724017]
19. Guo M, Ignatov M, Musier-Forsyth K, Schimmel P, Yang XL. *Proc Natl Acad Sci U S A*. 2008; 105:2331. [PubMed: 18272479]
20. Francin M, Kaminska M, Kerjan P, Mirande M. *J Biol Chem*. 2002; 277:1762. [PubMed: 11706011]
21. Francin M, Mirande M. *Biochemistry*. 2006; 45:10153. [PubMed: 16906773]
22. Monaco-Malbet S, Berthet-Colominas C, Novelli A, Battai N, Piga N, Cheynet V, Mallet F, Cusack S. *Structure*. 2000; 8:1069. [PubMed: 11080628]
23. Kingston RL, Vogt VM. *Mol Cell*. 2005; 17:166. [PubMed: 15664185]
24. Munro, AW.; Noble, MA. *Flavoprotein Protocols*. Vol. 131. Humana Press; Edinburgh: 1999.
25. Gray JJ. *Curr Opin Struct Biol*. 2006; 16:183. [PubMed: 16546374]
26. Hargittai MR, Mangla AT, Gorelick RJ, Musier-Forsyth K. *J Mol Biol*. 2001; 312:985. [PubMed: 11580244]
27. Gabor J, Cen S, Javanbakht H, Niu M, Kleiman L. *J Virol*. 2002; 76:9096. [PubMed: 12186893]
28. Sticht J, Humbert M, Findlow S, Bodem J, Muller B, Dietrich U, Werner J, Krausslich HG. *Nat Struct Mol Biol*. 2005; 12:671. [PubMed: 16041387]
29. Melamed D, Mark-Danieli M, Kenan-Eichler M, Kraus O, Castiel A, Laham N, Pupko T, Glaser F, Ben-Tal N, Bacharach E. *J Virol*. 2004; 78:9675. [PubMed: 15331700]
30. Francin M, Mirande M. *J Biol Chem*. 2003; 278:1472. [PubMed: 12417586]
31. Cusack S, Yaremchuk A, Tukalo M. *EMBO J*. 1996; 15:6321. [PubMed: 8947055]
32. Arnez JG, Moras D. *Trends Biochem Sci*. 1997; 22:211. [PubMed: 9204708]

Figure 1. Ab initio docking of Gag-CA-CTD domain onto human LysRS

A. Schematic of human LysRS and rigid-body docking results of CTD on human LysRS structure by HEX, ZDOCK and Dot. Each dot represents the core of docked CTD structures.

B. Docking of CTD on human LysRS in the refined space by AutoDock. The energy distribution of each cluster of solutions is shown below. **C.** Molecular dynamics refinement of the lowest energy model from AutoDock. After the MD refinement, CTD moved 3 Å towards the aminoacylation domain of LysRS and formed a tighter contact. The AutoDock solution of CTD and of the anticodon-binding domain of human LysRS, are shown in green. The MD final solution of CTD, and of anticodon-binding domain are shown in cyan and orange, respectively; the region 220–260 in the catalytic domain of human LysRS includes a hydrophobic patch (231–245) that overlaps with the characteristic motif-1 (238–260) in Class II AARSs. **D.** Surface representation of the final docking model. **E.** Closeup view of LysRS/CTD complex model. Helices 1–4 of CTD were colored from N-terminus (H1, red) to C-terminus (H4, green).

Figure 2. Model of human LysRS/tRNA^{Lys}/Gag-CA ternary complex
NTD of Gag capsid (Gag-CA) is shown in green. The N-terminal helix of human LysRS is shown as a surface representation. All Lys/Arg residues on the N-terminal helix are colored in deep blue, with most of them facing the bound tRNA.

Figure 3. Engineering human LysRS into a monomer

A. Human LysRS dimer is shown with one subunit as a schematic cartoon and the other as a surface representation. The aminoacylation and N-terminal anticodon-binding domains are colored the same as in Figure 1. **B.** Design of the mutagenesis on the peripheral region of the dimer interface. The predicted CTD binding site is located at the core of the LysRS dimer interface (colored in grey). The peripheral region of the dimer interface is shown in white, with the selected sites of mutations in black. The corresponding positions of these sites are also shown in the sequence scheme of LysRS. **C.** Flavodoxin insertion into human LysRS. *E. coli* Flavodoxin is inserted at the position of G310 on the peripheral region of the LysRS dimer interface. The linker is kept short to reduce the potential structural flexibility. A model of the fused *Flavo*-LysRS is shown with the flavodoxin structure in green. **D.** Overlay of Superdex 200 gel filtration profiles of human LysRS mutants. Mutated residues are shown in red, compared to the native residues shown in panel B. All proteins were loaded at a concentration range of 0.5–2 μ M. The vertical line gives the position of the LysRS dimer. Only *Flavo*-LysRS ran as a monomer. BSA is bovine serum albumin.

Figure 4. Characterization of monomeric *Flavo*-LysRS

A. *E. coli* flavodoxin and the purified *Flavo*-LysRS show a similar bright yellow color that indicates a properly folded flavodoxin domain in the fusion protein. **B.** Crosslinking assay of Thr385Cys mutants of both native LysRS and *Flavo*-LysRS, in the presence and absence of the DTT reducing agent, shows that *Flavo*-LysRS does not form a dimer in solution.

Figure 5. *Flavo*-LysRS retains Gag and tRNA^{Lys3} binding

A. StrepII-tag pull-down assay of binding to GagΔP6 that has been immobilized on Tactin beads through a D -desthiobiotin-elutable tag. **B –D.** EMSA assays of full-length LysRS and full-length *Flavo*-LysRS binding to *in vitro* transcribed human tRNA^{Lys3}. The binding affinity was calculated from the percentage of free/unbound tRNA on the gel.