Journal of Molecular Graphics and Modelling 71 (2017) 192–199

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Probing the stereospecificity of tyrosyl- and glutaminyl-tRNA synthetase with molecular dynamics

Karen Druart, Maeva Le Guennec, Zoltan Palmai, Thomas Simonson *

Department of Biology, Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, 91128 Palaiseau, France

ARTICLE INFO

Article history:
Received 13 September 2016
Received in revised form 8 November 2016
Accepted 11 November 2016
Available online 27 November 2016

Keywords:
Protein design
Ligand binding
Free energy
Computer simulations
Genetic code

ABSTRACT

The stereospecificity of aminoacyl-tRNA synthetases helps exclude D-amino acids from protein synthesis and could perhaps be engineered to allow controlled D-amino acylation of tRNA. We use molecular dynamics simulations to probe the stereospecificity of the class I tyrosyl- and glutaminyl-tRNA synthetases (TyrRS, GlnRS), including wildtype enzymes and three point mutants suggested by three different protein design methods. L/D binding free energy differences are obtained by alchemically and reversibly transforming the ligand from L to D in simulations of the protein–ligand complex. The D81Q mutation in *Escherichia coli* TyrRS is homologous to the D81R mutant shown earlier to have inverted stereospecificity. D81Q is predicted to lead to a rotated ligand backbone and an increased, not a decreased L-Tyr preference. The E36Q mutation in *Methanococcus jannaschii* TyrRS has a predicted L/D binding free energy difference $\Delta\Delta G$ of just 0.5 ± 0.9 kcal/mol, compared to 3.1 ± 0.8 kcal/mol for the wildtype enzyme (favoring L-Tyr). The ligand ammonium position is preserved in the D-Tyr complex, while the carboxylate is shifted. Wildtype GlnRS has a similar preference for L-glutaminyl adenylate; the R260Q mutant has an increased preference, even though Arg260 makes a large contribution to the wildtype $\Delta\Delta G$ value.

1. Introduction

Protein design has developed into an important tool for understanding and engineering protein structure and function [1–5]. An important advance has been the discovery of techniques to incorporate unnatural amino acids (UAAs) into proteins, both *in vitro* and *in vivo* [6–11]. This has expanded the amino set that can potentially be used in protein design [12,13]. The UAAs are inserted using an engineered aminoacyl-tRNA synthetase, which binds the UAA and attaches it to an engineered tRNA. The tRNA carries an anticodon that is complementary to a specific new codon, not used for ordinary amino acids, such as a rare STOP codon. The new codon and aaRS/tRNA pair can then be used to insert the UAA into proteins *in vivo*, effectively expanding the genetic code.

UAAs have been used to confer new properties, such as optical or spectroscopic activity for imaging or new types of chemical activity [10,11,14,15]. They can also be used, in principle, to alter protein structure. In particular, D-amino acids might be used to locally alter the protein backbone conformation [16–18]. Mixed L/D polypeptides have been produced *in vitro* [17–19]. It would also be of great interest to incorporate them into proteins *in vivo*, using an engineered aaRS/tRNA pair and an associated codon.

As a first step, several authors have characterized the stereospecificity of tyrosyl-tRNA synthetase (TyrRS) [20–26] and considered the problem of engineering it. Whereas most aaRSs have a strong preference for their L-amino acid substrate, TyrRS has a detectable, natural, tRNA-acylation activity for the D-tyrosine stereoisomer [20,21,23,26,25], being capable of charging D-Tyr onto $tRNA^{Tyr}$ to form D-Tyr-tRNA instead of the usual L-Tyr-tRNA. The production of D-Tyr-tRNA is important enough so that a D-aminoacyl-tRNA hydrolase enzyme has evolved to remove the D-amino acid and recycle the incorrectly charged tRNA's [27]. The natural TyrRS activity towards D-Tyr could in principle be increased and used for applications. Thus, First and coworkers [28,29] created a chimeric TyrRS enzyme by inserting the editing domain from PheRS into TyrRS; the editing domain hydrolyzes L-Tyr-tRNA but not D-Tyr-tRNA, leading to an increase in the proportion of D-Tyr-tRNA that is produced, and resulting in an inverted stereospecificity for the chimeric enzyme. However, the preference for D-Tyr is not very strong. Simonson et al used a mixture of automated protein design and *ad hoc* structure analysis to find a point mutant D81R of the *Escherichia coli* TyrRS with an inverted stereospecificity, having a strong preference for D-tyrosine over L-tyrosine [30]. However, its catalytic activity was weak. Thus, to effectively produce tRNA's

* Corresponding author. Fax: +33 169334860/4909.
E-mail address: thomas.simonson@polytechnique.fr (T. Simonson).

http://dx.doi.org/10.1016/j.jmgm.2016.11.007
1093-3263/

charged with a D-amino acid, there is still a need for improved variants of TyrRS or another aaRS.

Here, we use molecular dynamics simulations to characterize three additional TyrRS variants. We also consider another class I aaRS, glutaminyl-tRNA synthetase (GlnRS). Like TyrRS, GlnRS has a detectable tRNA-acylation activity for its D-amino acid substrate [21]. Its specificity determinants have been extensively studied [31–33]. We use alchemical free energy simulations to compute relative L/D binding free energies [34,35]. A similar free energy simulation method was used to explain the stereospecificities of aspartyl-tRNA synthetase [36] and of the threonyl-tRNA synthetase editing domain [37]. It was also applied earlier to some of our candidate *E. coli* TyrRS variants, including D81R [30]. Here, we first consider another *E. coli* TyrRS variant, D81Q, which was suggested by automated protein design [30]. Next, we consider wildtype *Methanococcus jannaschii* TyrRS and one of its mutants. The mutation, E36Q, is homologous to the D41N mutation of the *E. coli* enzyme, a candidate position identified earlier [30] based on a simple analysis of the TyrRS structure. Finally, we considered the wildtype GlnRS and one of its mutants: the R260Q point mutant. This mutation was suggested by a component analysis of the wildtype binding free energies, described here. GlnRS is one of the aaRSs that requires tRNA for amino acid activation. Therefore, for GlnRS, we considered its complexes with the glutaminyl-adenylate substrate (GlnAMP), instead of the amino acid Gln. We compared the binding free energies of L- and D-glutaminyl-adenylate in the presence of the tRNA co-substrate.

We obtain the following main results. First, for the D81Q *E. coli* TyrRS variant, molecular dynamics simulations predict that the L-Tyr and D-Tyr ligand backbones assume an orientation that is incompatible with a productive enzymatic reaction. The stereospecificity of this variant is predicted to be slightly larger than that of the wildtype *E. coli* TyrRS, and similar to the homologous D81N variant studied earlier [30]. Second, for the archaeal *M. jannaschii* enzyme, our simulations predict a stereospecificity that is similar to yeast TyrRS, and slightly larger than that of the *E. coli* enzyme. A point mutant of the *M. jannaschii* enzyme, E36Q, is predicted to have a reduced stereospecificity, though not an inverted one. Finally, for *E. coli* GlnRS, our simulations predict about the same level of stereospecificity as for yeast TyrRS. The R260Q point mutant GlnRS has a slightly increased L-GlnAMP preference. Overall, among the two TyrRS mutants and the GlnRS mutant, the simulations do not reveal any inverted stereospecificities. They do show good consistency with earlier simulations and experiments, and they provide useful information regarding mutant structures, the sources of specificity, and the protein design strategies employed.

2. Materials and methods

2.1. Molecular dynamics models for TyrRS

For the D81Q variant of *E. coli* TyrRS, the MD setup was the same one used earlier for wildtype and several TyrRS variants [30]. Starting structures were generated from the crystal structure of *E. coli* TyrRS bound to a tyrosyl adenylate analog (PDB code 1VBM) [38]. The protein is a symmetric homodimer; we focus on one of the ligands and binding sites. The ligand was truncated to give L-Tyr. Sidechain protonation states were assigned by visual inspection and calculations with the PropKa program [39,40], applied to six *E. coli* TyrRS X-ray structures. The only unusual case was the Asp182 sidechain, modeled in a protonated state. For five of the six X-ray structures, PropKa predicted a neutral protonation state in the presence of bound Tyr. This protonation state is also suggested by visual analysis of the structures. Indeed, the Asp182 sidechain is completely buried when the amino acid ligand is bound, and can donate a hydrogen bond to the ligand sidechain hydroxyl. A buried water molecule close to the Tyr sidechain is conserved in several crystal structures, including 1VBM, and was included in the model. The mutant Q81 sidechain was placed using the Scwrl4 sidechain reconstruction program [41], which selects the optimal Q81 rotamer, with the surrounding sidechains in their wildtype, X-ray orientation. The initial Q81 orientation pointed away from the ligand, but during the molecular dynamics equilibration phase, it rotated so as to coordinate the ligand ammonium group with its sidechain carbonyl, similar to the wildtype Asp81 sidechain. This coordination was mostly preserved throughout our MD simulations even though Gln81 is partly solvent exposed and mobile. For D-Tyr placement, we started by superimposing D-Tyr onto the X-ray L-Tyr, choosing a D-Tyr rotamer that led to a very good superposition, then performing 100 steps of energy minimization. Protein residues with no nonhydrogen atoms within 28 Å of the ligand center were then deleted; protein atoms from 20 to 28 Å from the ligand center were held in place by weak harmonic restraints during the molecular dynamics simulations (MD), to mimic the structural effect of the deleted groups [42]. A cubic box of water with an 80 Å edge was overlaid and waters overlapping protein or ligand were removed (except the one buried water; see above). Three chloride ions were included to ensure electroneutrality. The overall model contained about 49,000 atoms; the same cutoff was used for all the other systems, described below, leading to very similar total atom numbers. Notice that the system is truncated for reasons of efficiency, and because we expect that the interactions that account for stereospecificity are unlikely to act significantly over distances of more than 15–20 Å. Thus, we do not expect that a larger simulation model would give significantly better accuracy.

MD was performed at room temperature and pressure, using a Nose-Hoover thermostat and barostat. Long-range electrostatic interactions were treated with a Particle Mesh Ewald approach and tnfoil boundary conditions [43,44]. The CHARMM27 forcefield [45] was used for the protein and ions and a slightly modified TIP3P model [46] for water. For the Tyr ligand, CHARMM27 parameters were used, except for the backbone charges, which are not available in the standard forcefield. Because of this, charge values were derived in our recent work [30] by an *ab initio* quantum chemistry method typically used for the Amber ff99SB force field [47,48]. In principle, we could have applied the standard parameterization approach used to develop the Charmm force field [45]. However, this procedure is much more expensive and complex, since it involves repeated *ab initio* calculations on several water-solute complexes, done in an iterative manner. Given the high quality of the Amber force field, we felt the corresponding parameterization approach would provide a good approximation. The values (in units of a proton charge) are as follows: C_α: 0.0204, H_α: 0.0741, C_β: −0.1172, H_β's: 0.0945; 0.7538 and −0.7051 for the backbone carboxylate carbon and oxygens; −0.3025 and +0.2770 for the backbone ammonium N and hydrogens. Simulations were done with the Charmm and NAMD programs [49,50].

For the wildtype and E36Q *M. jannaschii* TyrRS, the 1J1U X-ray structure was used [51], which contains L-Tyr and tRNA as ligands. The tRNA was deleted. Asp158 (homologous to *E. coli* Asp182) was modeled in its protonated state, so as to donate a hydrogen bond to the L-Tyr hydroxyl (which then donates a hydrogen bond to Tyr32). The same buried water was included as above. For the E36Q mutant, Gln was oriented with its sidechain oxygen towards the Tyr ligand. D-Tyr was positioned as above. The corresponding models were then solvated, truncated, equilibrated, and simulated as above.

2.2. Molecular dynamics models for GlnRS

The starting structure was a complex (PDB code 1QTQ) between *E. coli* GlnRS, tRNA, and the GlnAMP analog 5′-O-(L-glutaminyl-sulfamoyl) adenosine, or QSI [52]. After attempting simulations