The Discovery of *E. coli* Methionyl-tRNA Synthetase Mutants That Activate Azidonorleucine at Low Concentrations: Collecting *In Vivo* Data for an *In Silico* Screen

I. Caglar Tanrikulu (a), David A. Tirrell (b), William A. Goddard III (c)

(a) Biochemistry and Molecular Biophysics Option,  
(b) Division of Chemistry and Chemical Engineering,  
(c) Materials and Process Simulation Center,  
California Institute of Technology, Pasadena, CA 91125, USA

*In vivo* incorporation of novel non-natural amino acids (*nnaas*) allows the introduction of various functionalities into proteins that are not represented in the canonical set of twenty, while maintaining high protein yields that many applications require. The incorporation is largely controlled by a class of high fidelity enzymes called aminoacyl-tRNA synthetases (*aaRSs*), which are responsible for charging a tRNA with its cognate amino acid. Even though incorporation of *nnaa* is possible through natural *aaRSs*, many *nnaas* are substantially different in size and shape compared with their natural counterparts, and are not good substrates for natural *aaRSs*. In such cases, the incorporation is achieved by modifying the activity and/or the specificity of a natural *aaRS*, so that it can charge its cognate tRNA with a *nnaa* in amounts needed to sustain the protein synthesis at a high level.

One such *nnaa* is the methionine analog, azidonorleucine (*Anl*). *Anl* incorporation into proteins is not detectable by mass spectrometry in the presence of wild-type *E. coli* methionyl-tRNA synthetase (*MetRS*). By randomizing four positions (L13, P257, Y260, H301) in the *MetRS* gene, Link and co-workers (*Proc. Natl. Acad. Sci. USA* 2006. 103, 10180) have identified a series of *MetRS* mutants through library screening that allow about 50% incorporation of *Anl* into methionine positions in the presence of 8 mM *Anl*. In addition, they were able to identify a single mutant (L13G) through rational design that shows a higher incorporation level (95%) at lower *Anl* concentrations (1 mM).

By evaluating the binding energies of the experimentally characterized mutants, we were able to distinguish between *MetRS* variants active and inactive towards *Anl* successfully. Based on this result, we evaluated the binding energies of a computational *MetRS* library, where three (L13, P257, Y260) of the original four residues in the screen were randomized. Through this computational screen we were able to retrieve back the L13G mutant as one of the top ranking candidates. However, experimental testing of other top ranking *MetRS* mutants failed to reveal any species that match the activity of the L13G variant towards *Anl*.

In order to tune our computational parameters to reveal functional mutants, it was important to work with a broad set of mutation data focused on a few positions. For this purpose, we employed an experimental screening strategy to identify functional mutants from a *MetRS* library randomized at three positions (L13, Y260, H301). A library of this size can easily be explored with complete coverage both experimentally (32³=32768 variants) and computationally (20³=8000 mutants), and the result of these screens can later be compared.
The experimental library was constructed with a >99% coverage of the library diversity. This library was explored through a high-throughput screen established by Link et al. In this screen, the cells harboring active MetRS mutants display azide groups on their surfaces, which can in turn be fluorescently labeled. Fluorescent cells can then be enriched using a fluorescence-activated cell sorter (FACS). Preliminary data from our screens has exposed a diverse set of mutants that match, or exceed the activity of the L13G mutant.

We are currently involved with: 1) identifying a set of MetRS triple-mutants that can activate Anl at very low concentrations, and 2) identifying a second set of mutants that can only activate Anl at relatively higher concentrations. Using two such distinct sets of mutants, we intend to test the effectiveness of our computational screen and improve on our methods, so that MetRS mutants active on novel mnaas can be identified solely based on a computational screen.

![Median Population Fluorescence Intensity at 1.0 mM Anl](image)

Median fluorescence of labeled cells harboring various MetRS variants obtained from FACS: Each mutant’s name refers to the identity of the residues it has at positions 13, 260, and 301 in MetRS. The NLL (L13N-Y260L-H301L) mutant exhibits the highest fluorescence at 1.0 mM Anl, as well as at 0.3 mM Anl.