Development of a high-throughput assay to identify possible foldable *de novo* proteins

Background: Different species have different proteins encoded in their genome. It has been thought that proteins evolve via gene duplication, but recently published work suggested, that there are proteins, called *de novo*, that don't show any similarity to other genes, meaning that they can't originate from duplication. The expression of the *de novo* proteins is still very difficult. Therefore, an assay which could help to test *de novo* protein folding in high-throughput would help to act as valuable starting point.

Objectives: In this project the different enzyme restriction sites (XhoI, HindIII, etc.) need to be implemented into the pSALect wildtype plasmid via PCR. In a next step this engineered plasmid will be self-ligated and cloned into E.cloni (*E.coli)* cells. After sequencing to assure the correctness of the engineered plasmid, different kinds of *de novo* genes should be cloned into the pSALect vector using the new introduced restriction sites. In a follow up assay the possible folding of these *de novo* gene constructs should be controlled via ampicillin agar plates. The idea is that only cells, expressing the *de novo* protein ampicillin construct will grow.

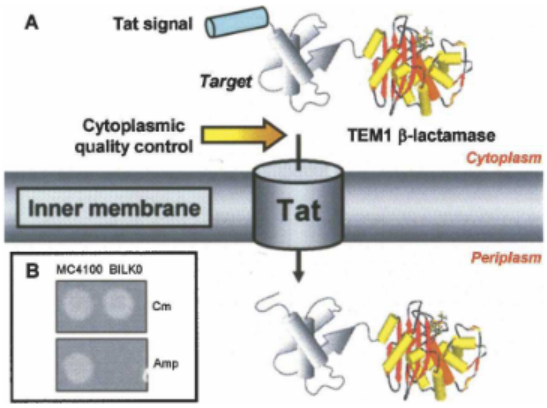

Figure 1: Exploiting the Tat pathway's folding quality control feature for monitoring protein solubility.

Requirements:

- Interest in evolution at the level of individual proteins
- Interest in lab work on DNA and protein level and basic knowledge of PCR, DNA-cloning, protein expression & purification

Methods:

- Molecular biochemistry (DNA-cloning, PCR, restriction digest, ligase, etc.) & expression and purification of proteins
- Biochemical characterization of the proteins via SDS-gels

Supervision: Dr. Andreas Lange, Room 205, andreas.lange@uni-muenster.de and Margaux Aubel, Room 19, m.aubel@uni-muenster.de, Molecular Evolution and Bioinformatics Group (http://bornberglab.org/).

Selected Literature:

1 Adam C. Fisher, Wookin Kim, and Matthew P. Delisa
Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway, *Protein Science*, **2006**
2 Keeling, DM. , Garza, P., Nartey, CM., Carvunis, AR.
Philosophy of Biology: The meanings of 'function' in biology and the problematic case of de novo gene emergence
elife, **2019**
3 Erich Bornberg-Bauer, Klara Hlouchova, and Andreas Lange
Structure and function of naturally evolved de novo proteins
COSB, **2021**
4 Andreas Lange, Prajal H Patel, Brennen Heames, Adam M Damry, Thorsten Saenger, Colin J Jackson, Geoffrey D Findlay, Erich Bornberg-Bauer
Structural and functional characterization of a putative de novo gene in Drosophila
Nat. Comms, **2021**