

NanoDSF sample preparation to enable high quality data for machine learning models

Lovisa Brandt,^a

^a EnginZyme AB,

Tomtebodavägen 6, 171 65 Solna, Sweden

lovisa@enginzyme.com

Nano differential scanning fluorimetry (nanoDSF) is a technique which allows for analysis of protein unfolding events in response to a gradual increase in temperature. [1] Comparing the proteins' thermal unfolding provides an important dataset when improving protein stability. For accurate analysis and comparison between datasets, reproducibility and high-quality data are two very important factors. Therefore, it is necessary to assess and establish which parameters are important to achieve reproducible data and what their boundaries are.

The goal of this project was to investigate the optimal parameters for measurements to be reproducible and of high quality for the purpose of training machine learning models for improving protein stability. To train the models a dataset of at least 200 variants is normally screened. NanoDSF measurements can be impacted by difference in protein purity and different additives in the protein environment [2]. Results from this project showed how limiting additives and impurities in the protein samples enables reproducible data for protein's thermal unfolding events.

The findings in this project allowed the workflow to facilitate high quality data from the screening of 200 variants in 2 weeks, starting from plasmids. Furthermore, a panel of a total of 400 variants was able to be screened and produce a high-quality dataset for machine learning models.

[1] Kim SH, Yoo HJ, Park EJ, Na DH. *Pharmaceuticals (Basel)*. **2021**, *15*, 29.

[2] DIN SPEC 13290. *Protein analysis - Evaluation criteria for thermal unfolding methods (DSF and DSC); Text in German and English*. **2021-09**.