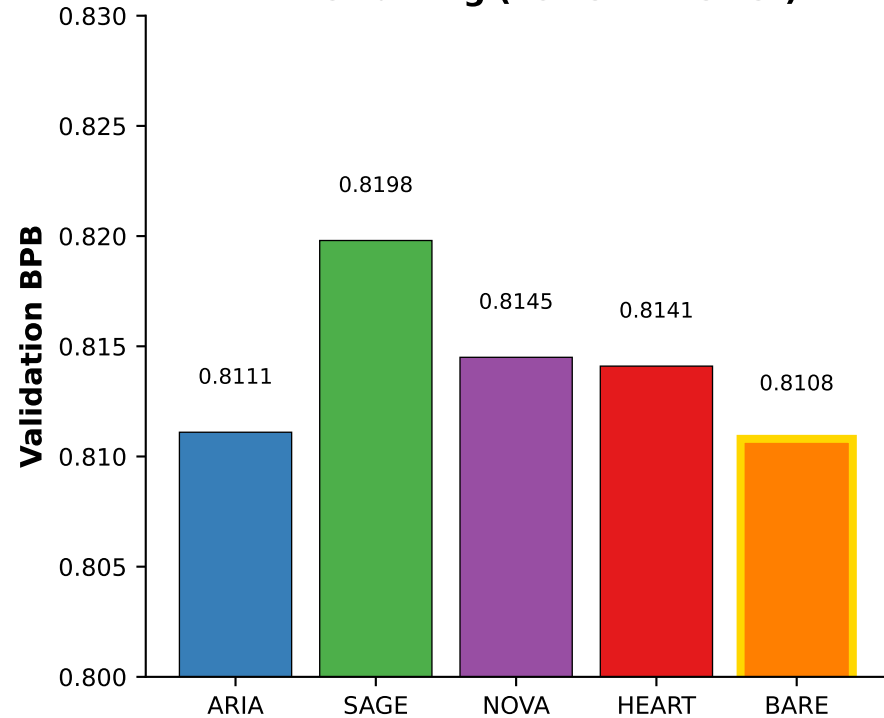
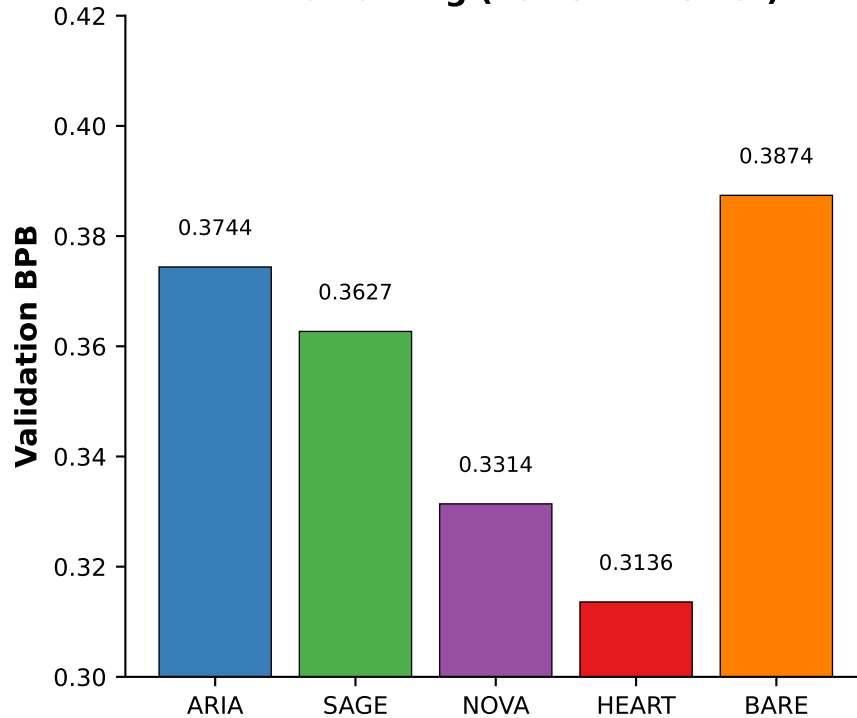


# Training Dynamics Across Stages

## A. Pretraining (Lower = Better)



## B. Midtraining (Lower = Better)



## C. SFT Training Loss

