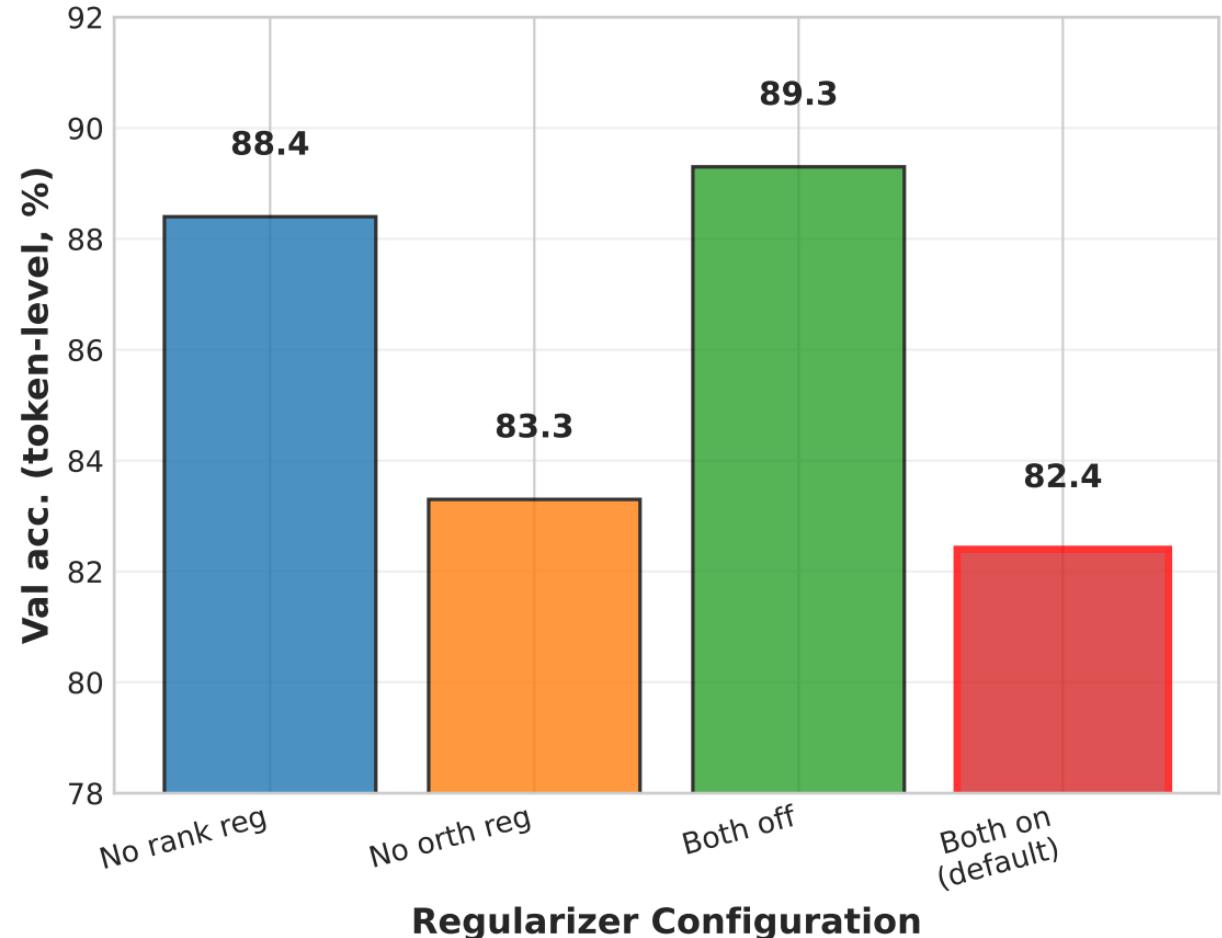


Token-Level Validation Accuracy vs. Regularization



Train Loss vs. Regularization

