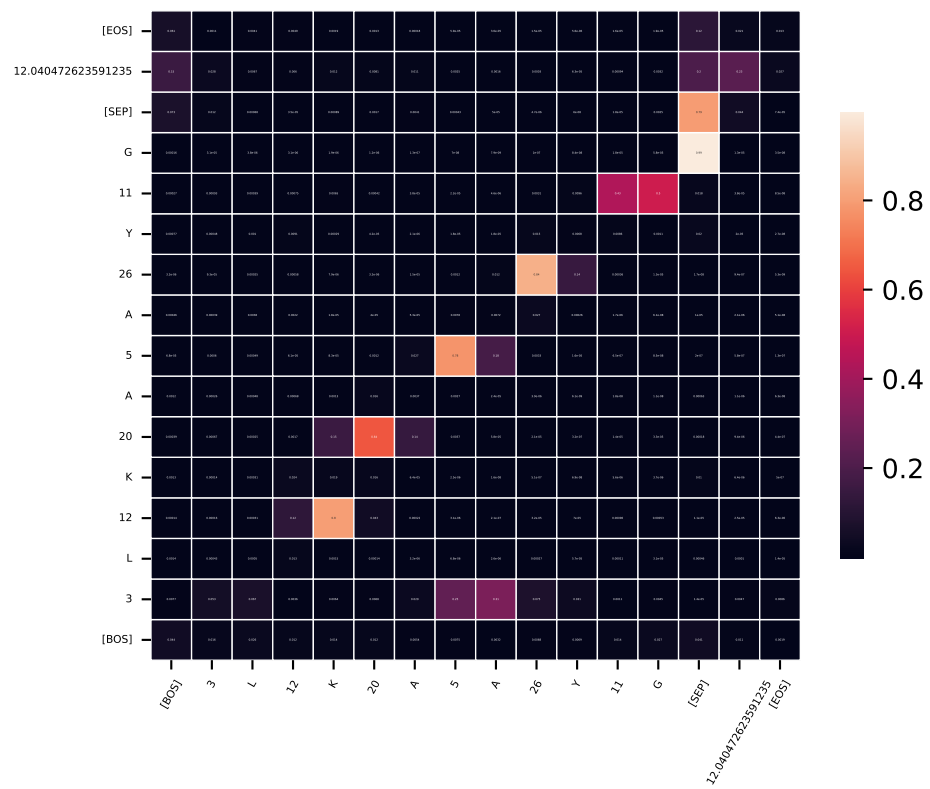
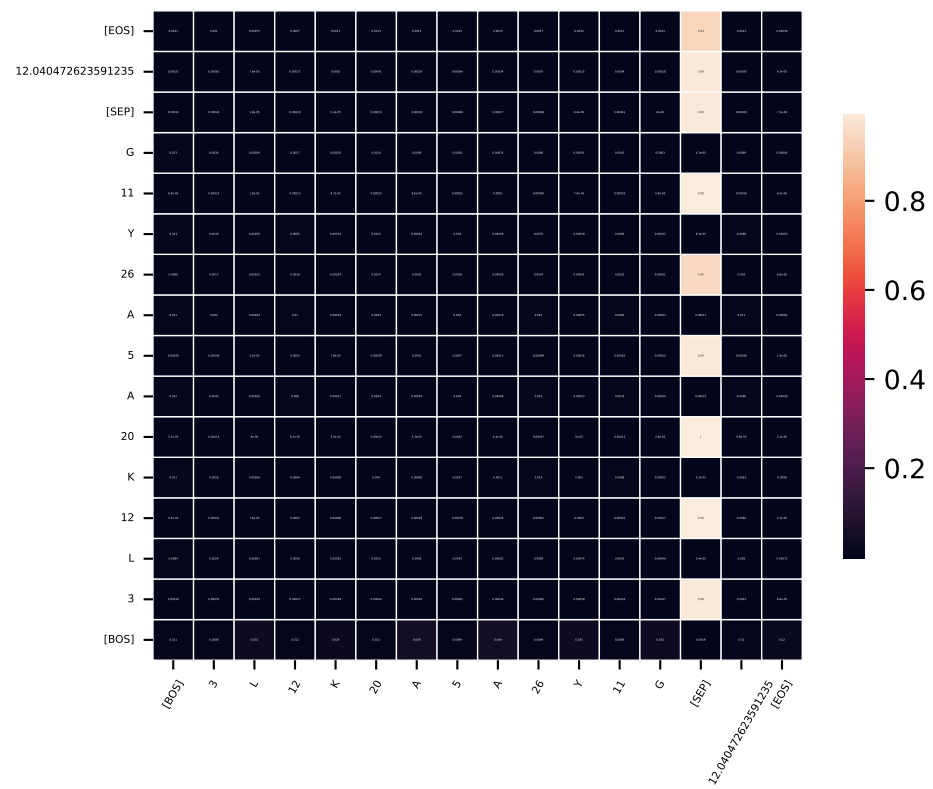
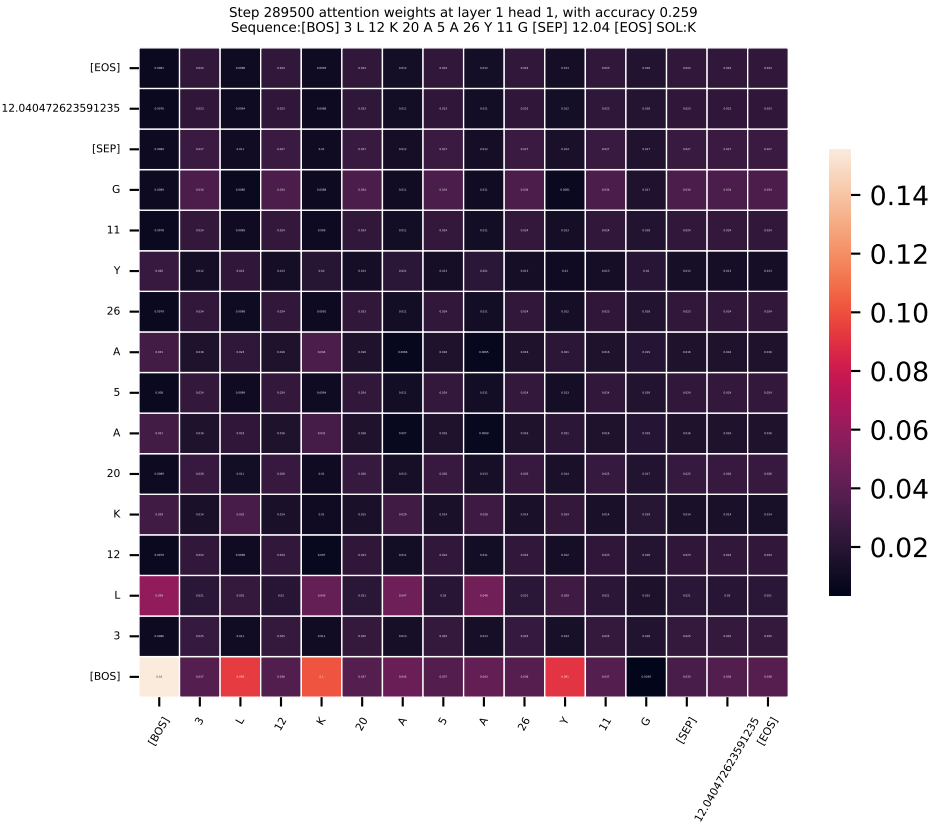
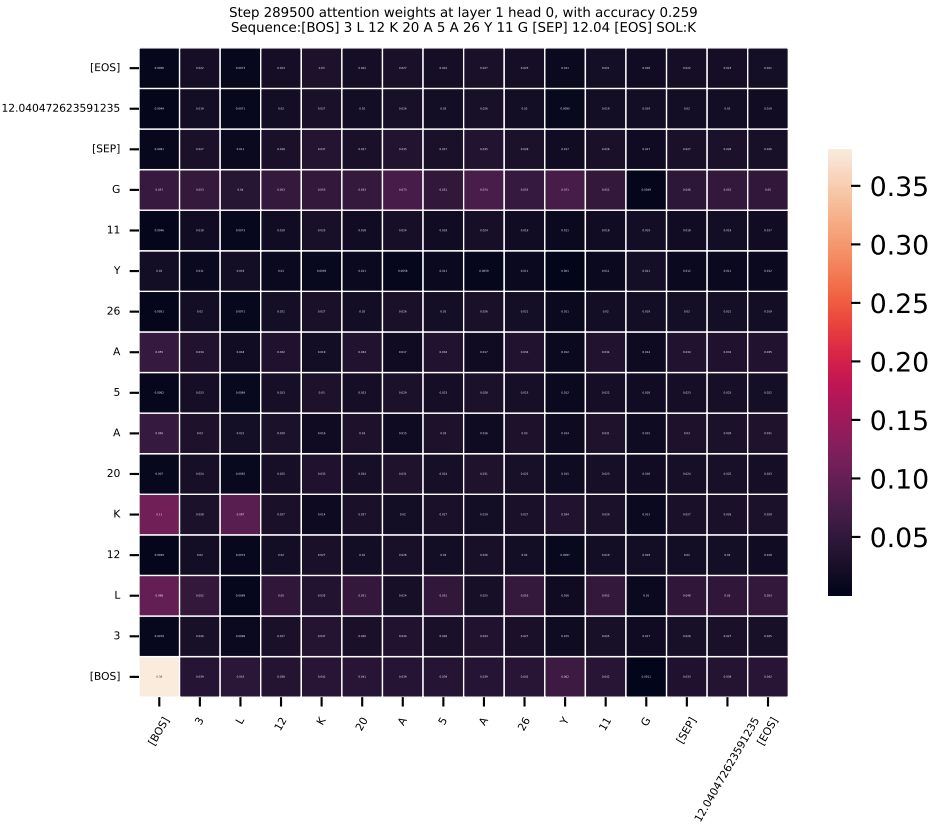


Step 289500 attention weights at layer 0 head 0, with accuracy 0.259
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K

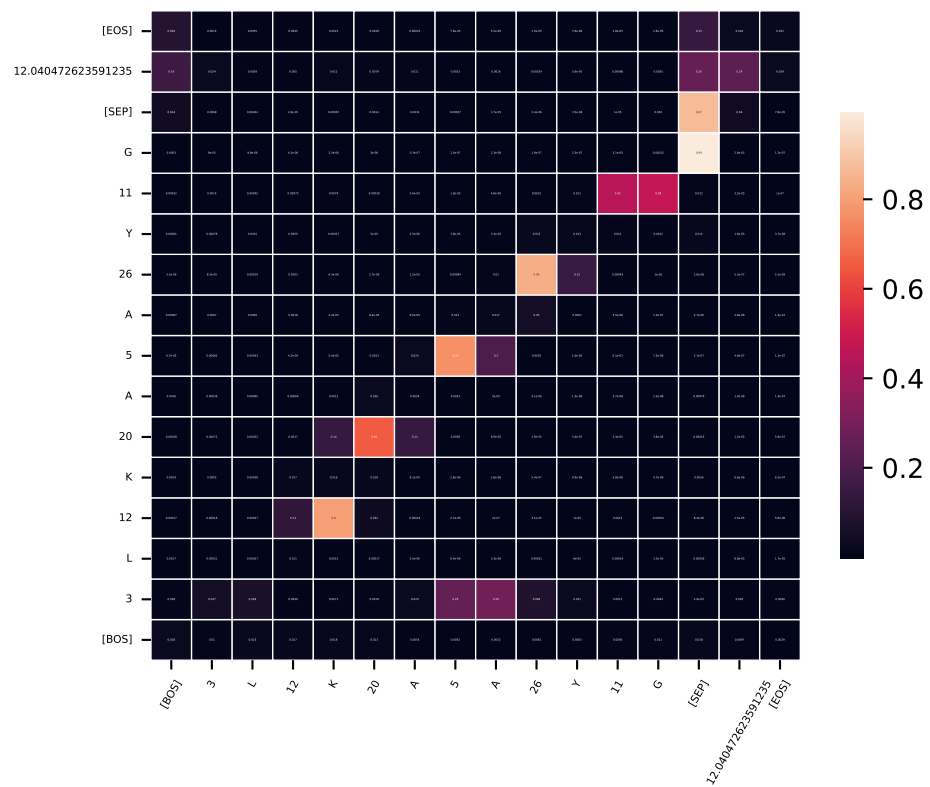


Step 289500 attention weights at layer 0 head 1, with accuracy 0.259
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K

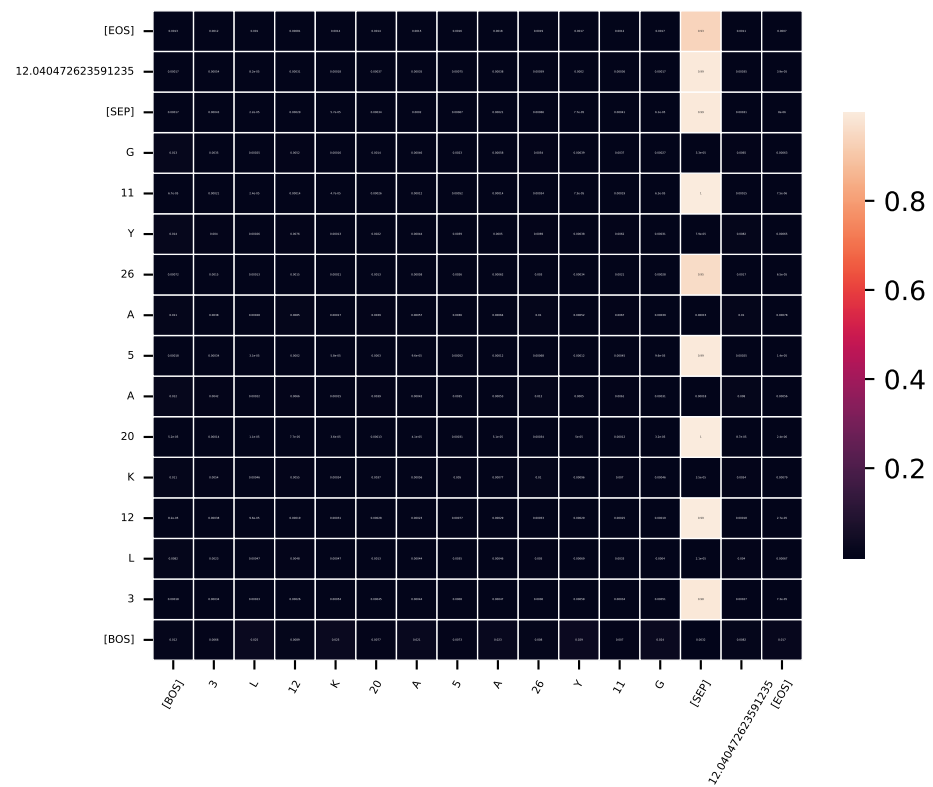


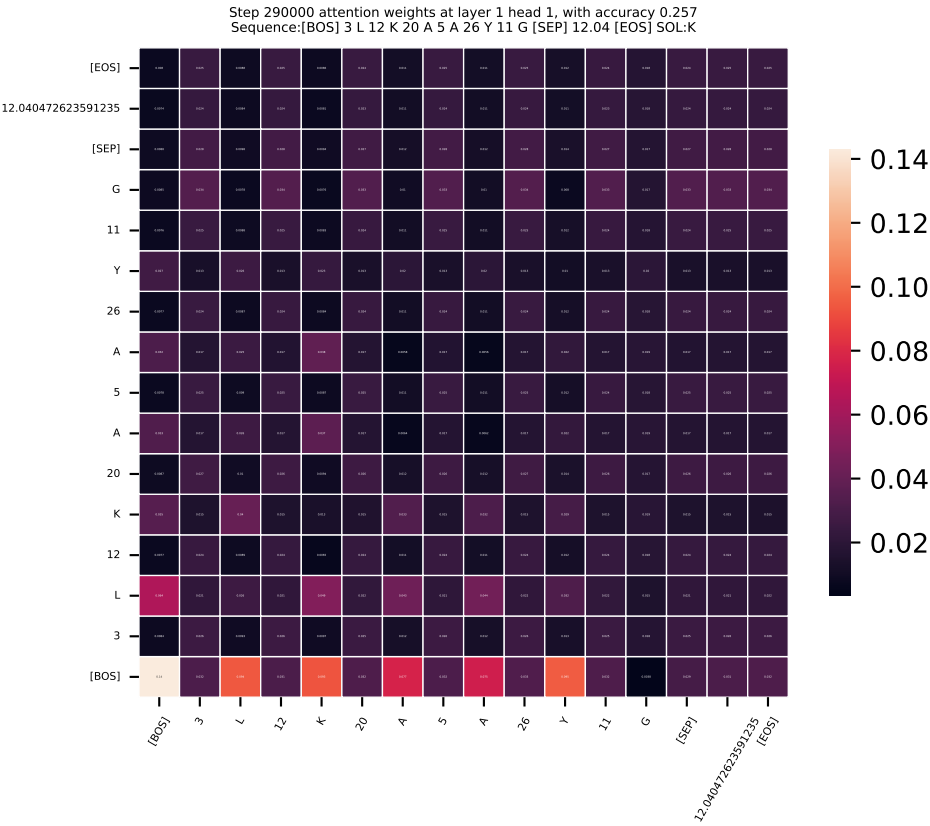
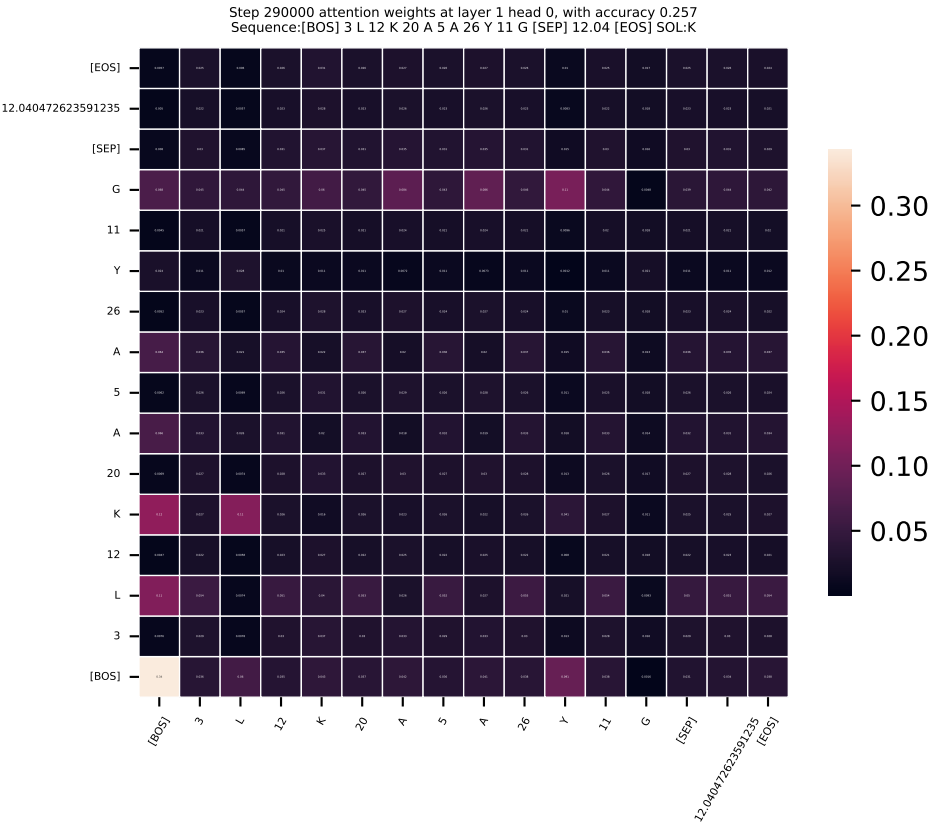


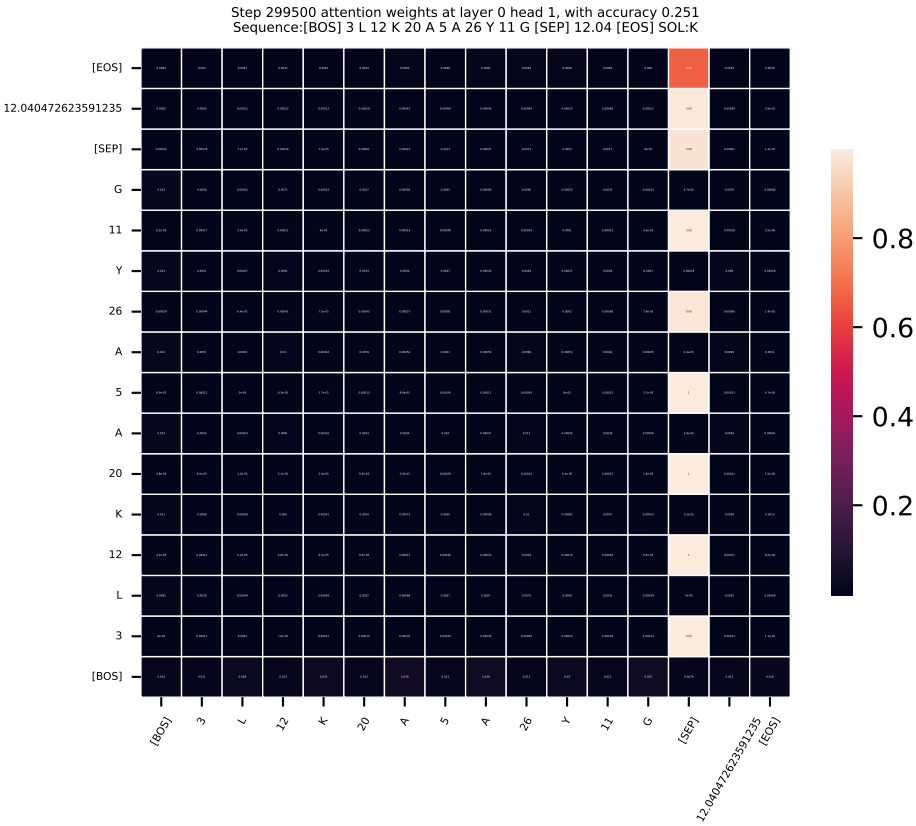
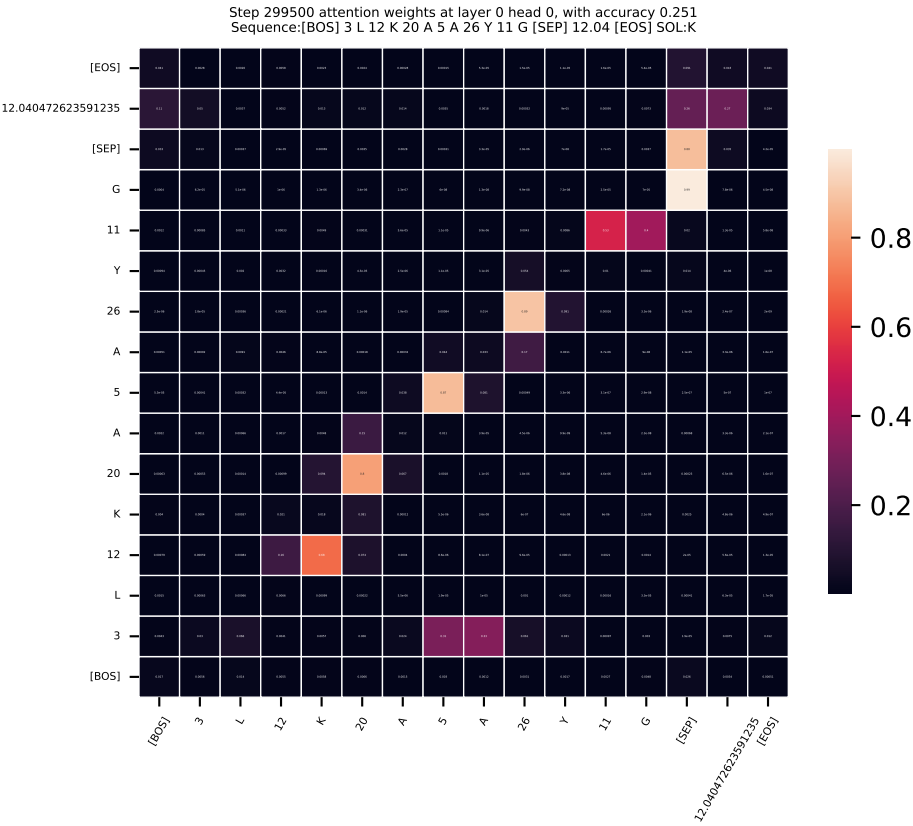
Step 290000 attention weights at layer 0 head 0, with accuracy 0.257
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K

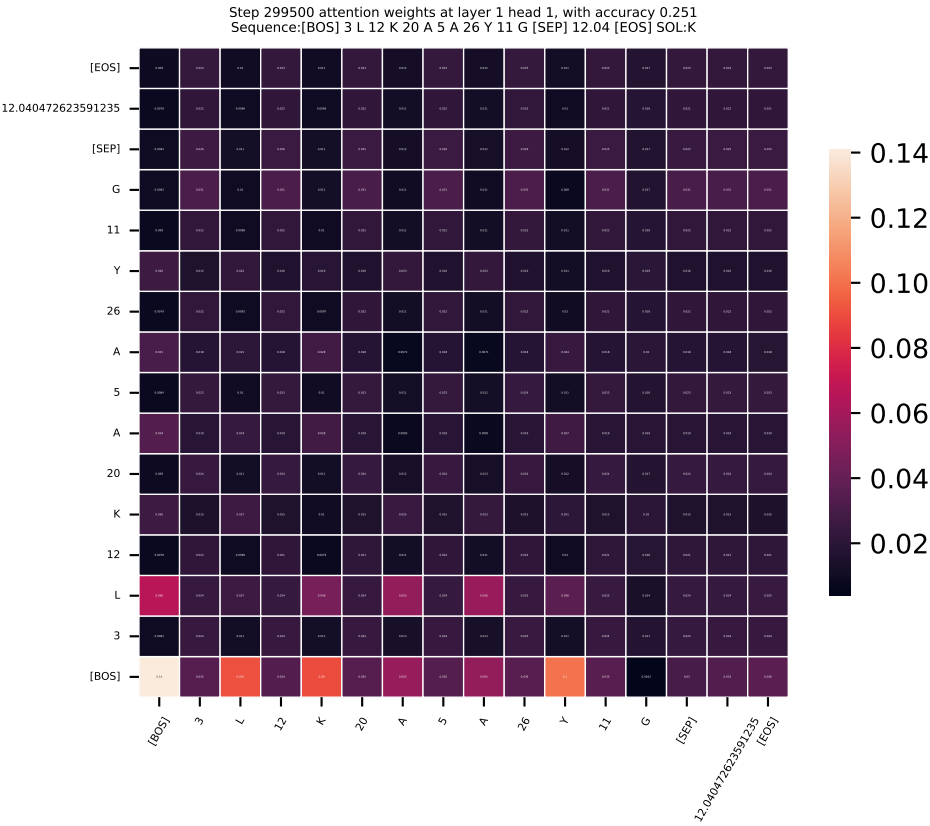
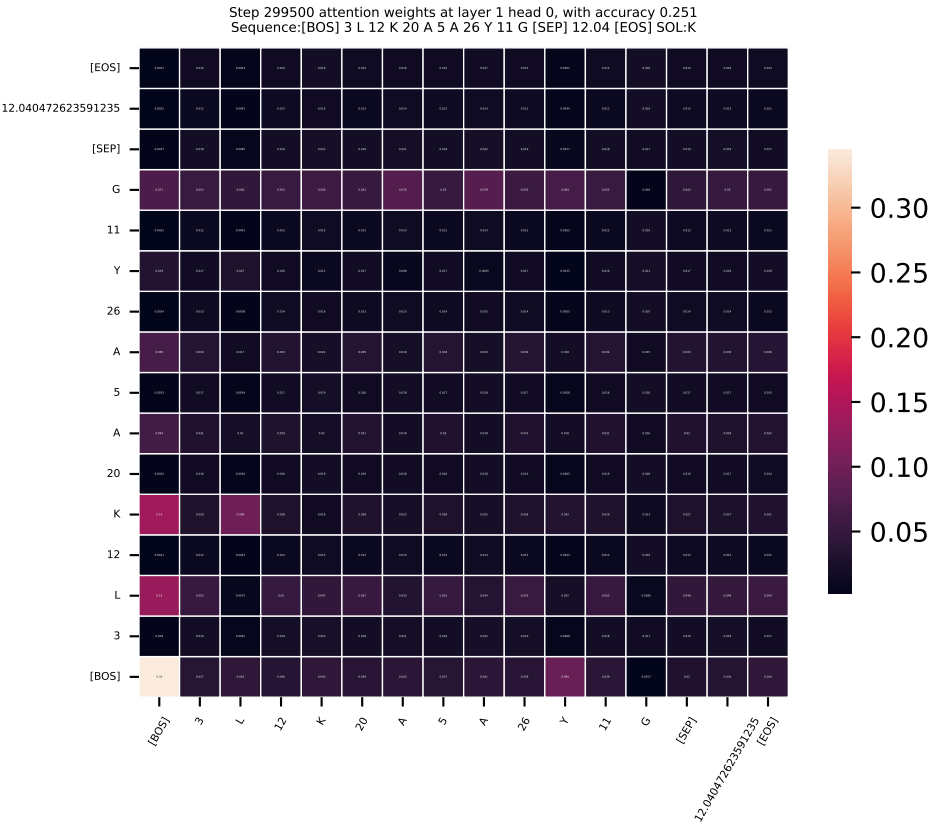


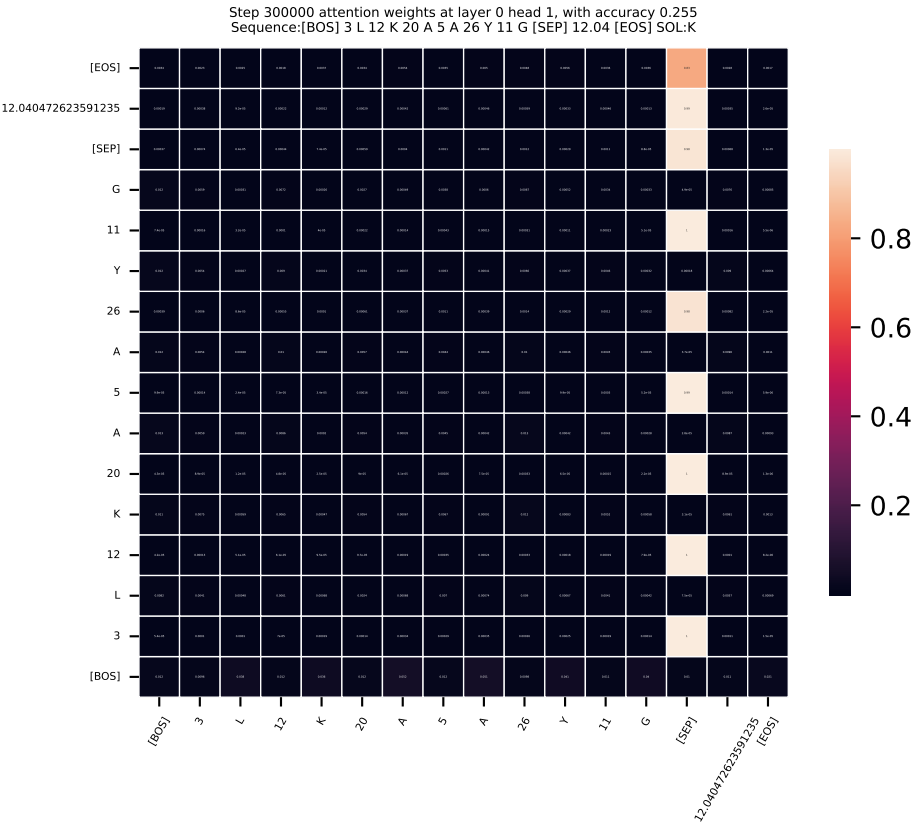
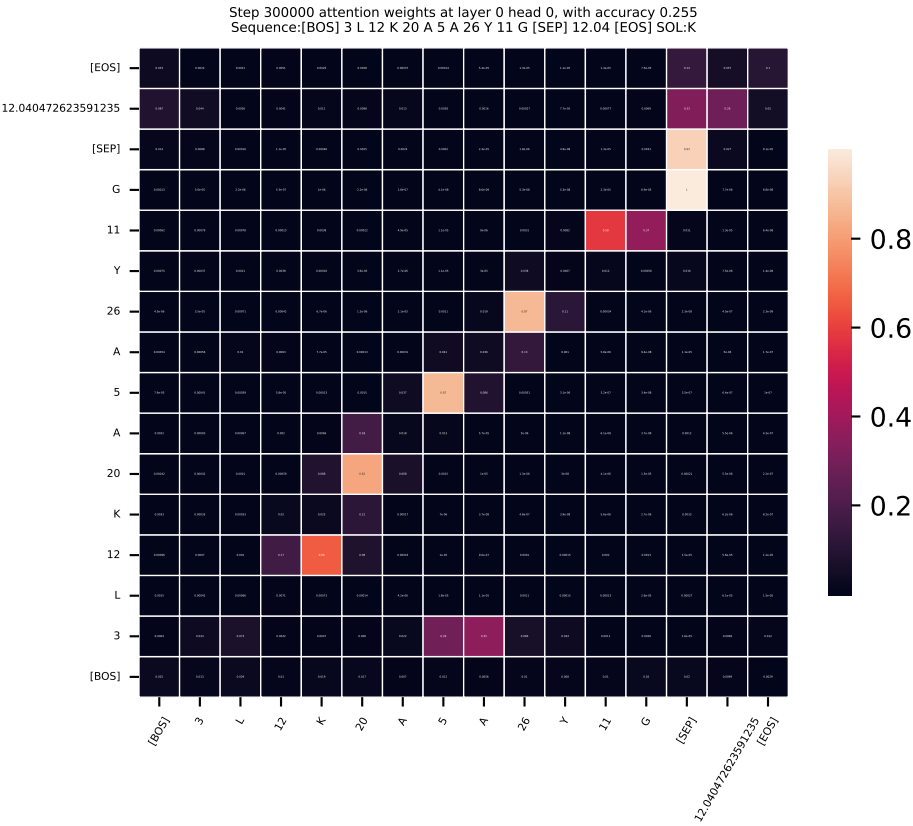
Step 290000 attention weights at layer 0 head 1, with accuracy 0.257
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K

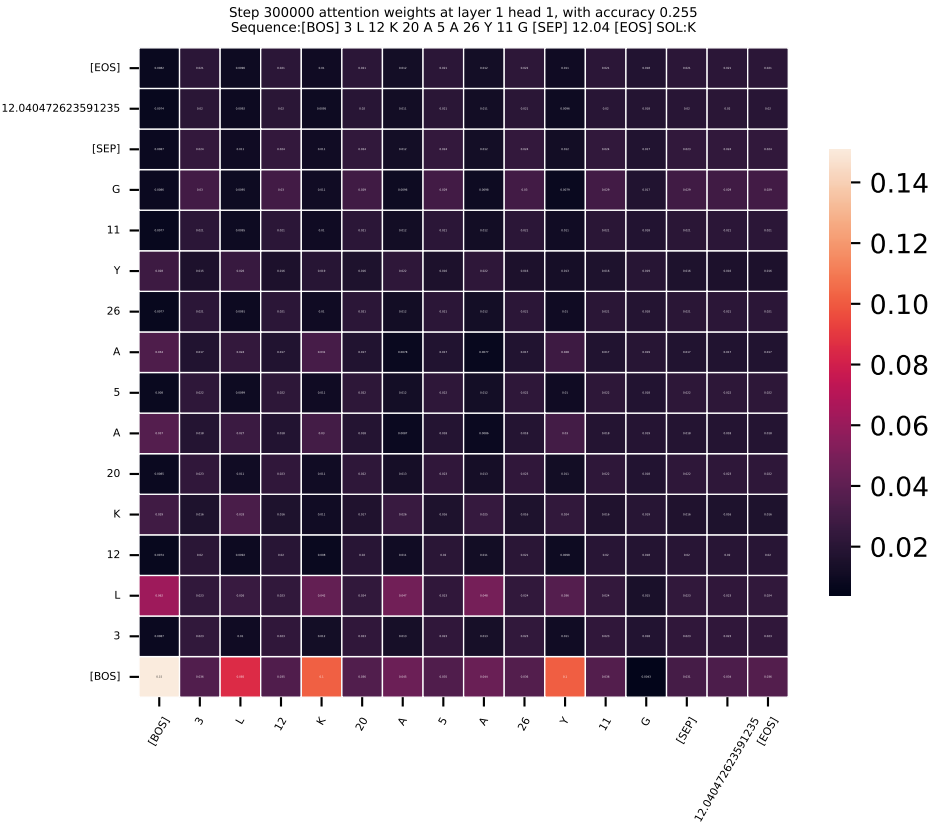
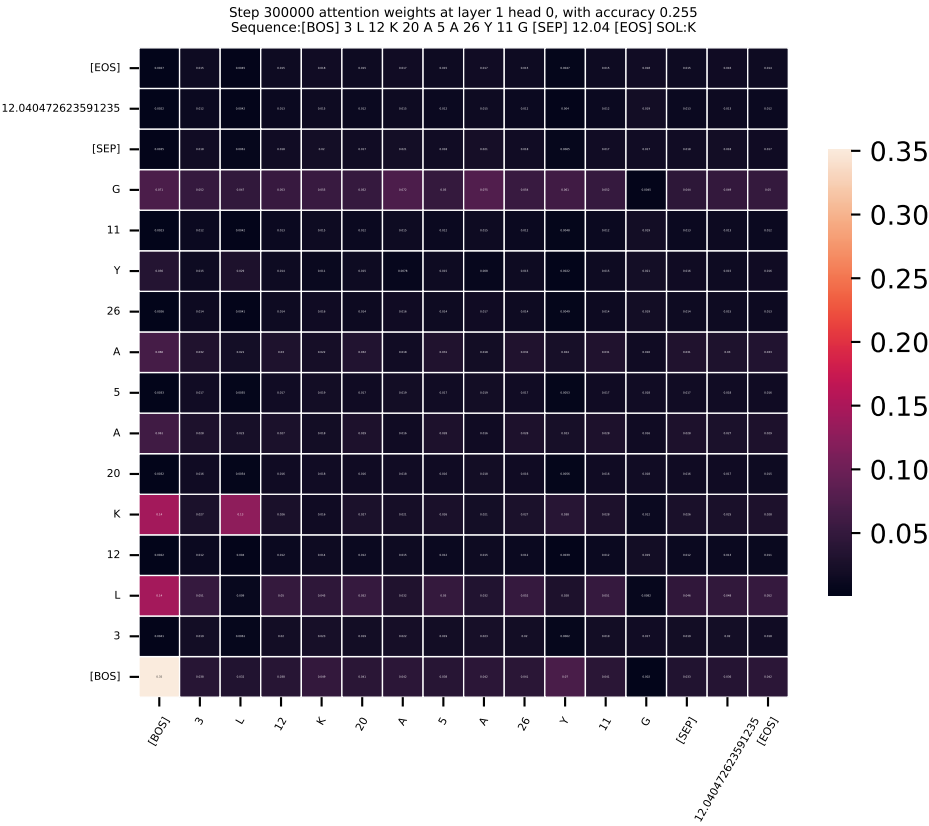


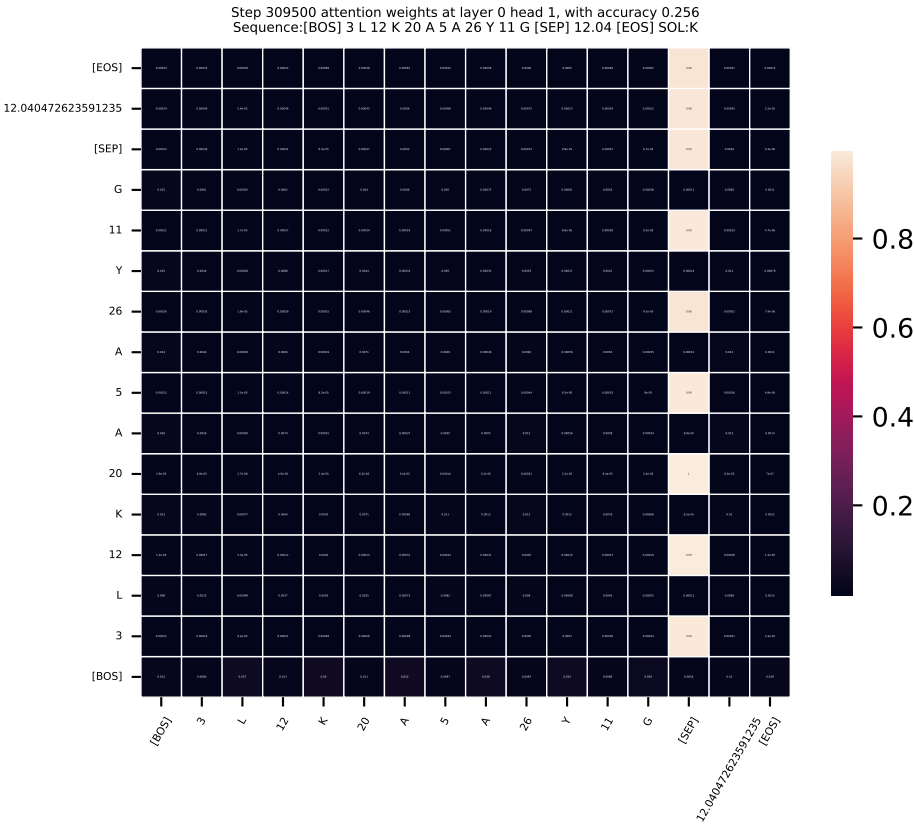
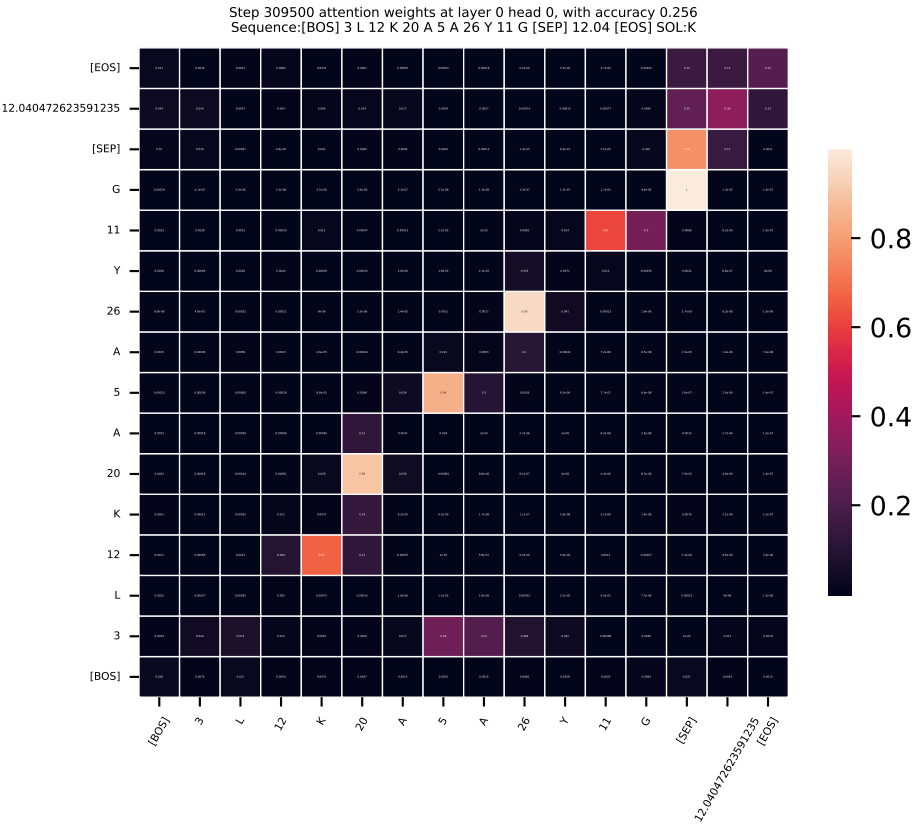


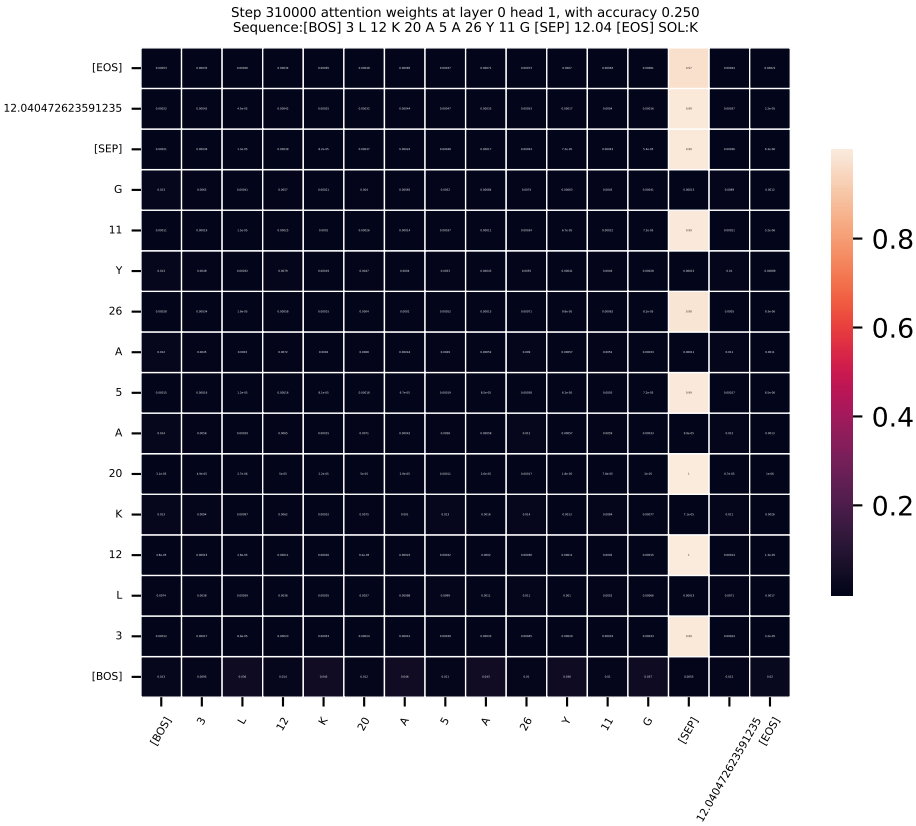
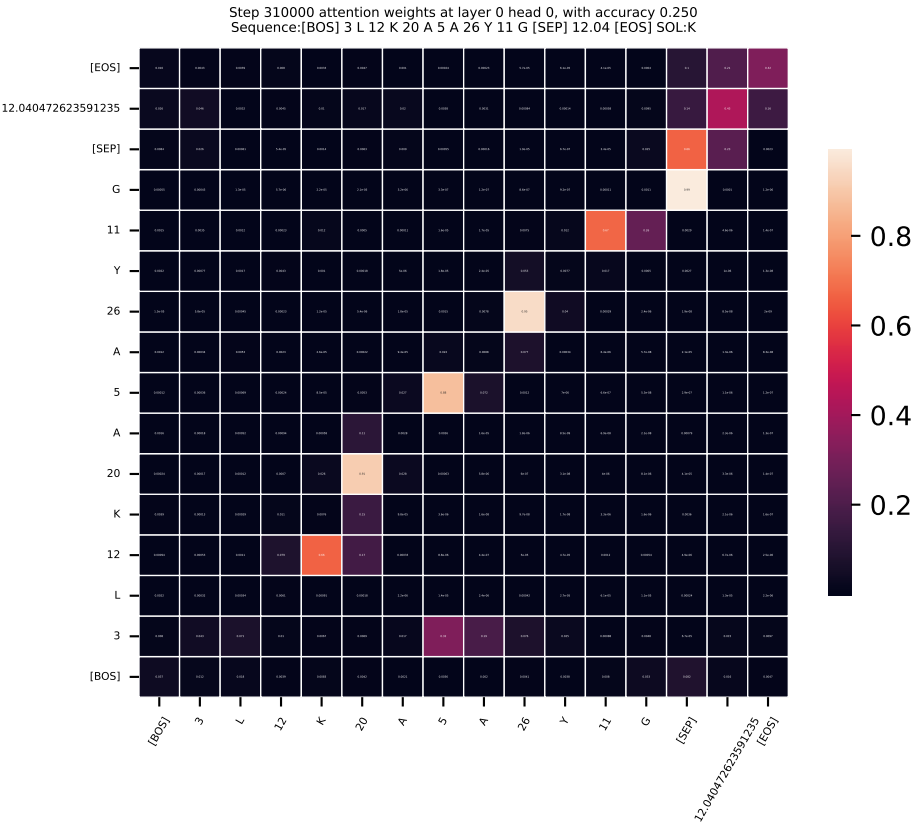


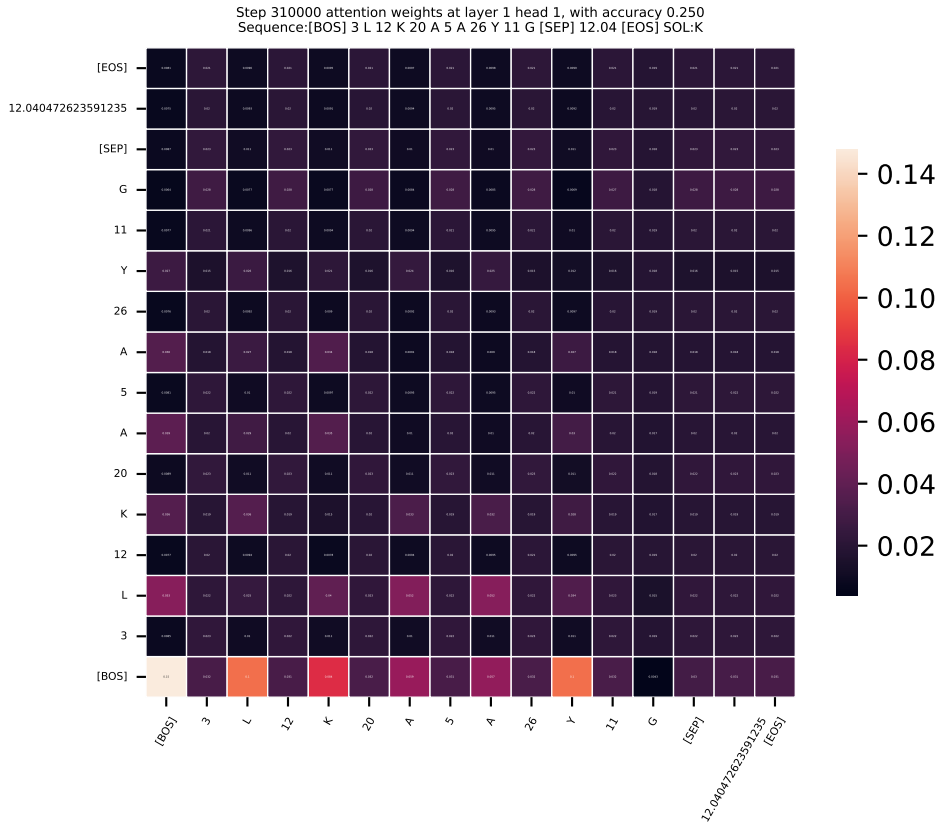
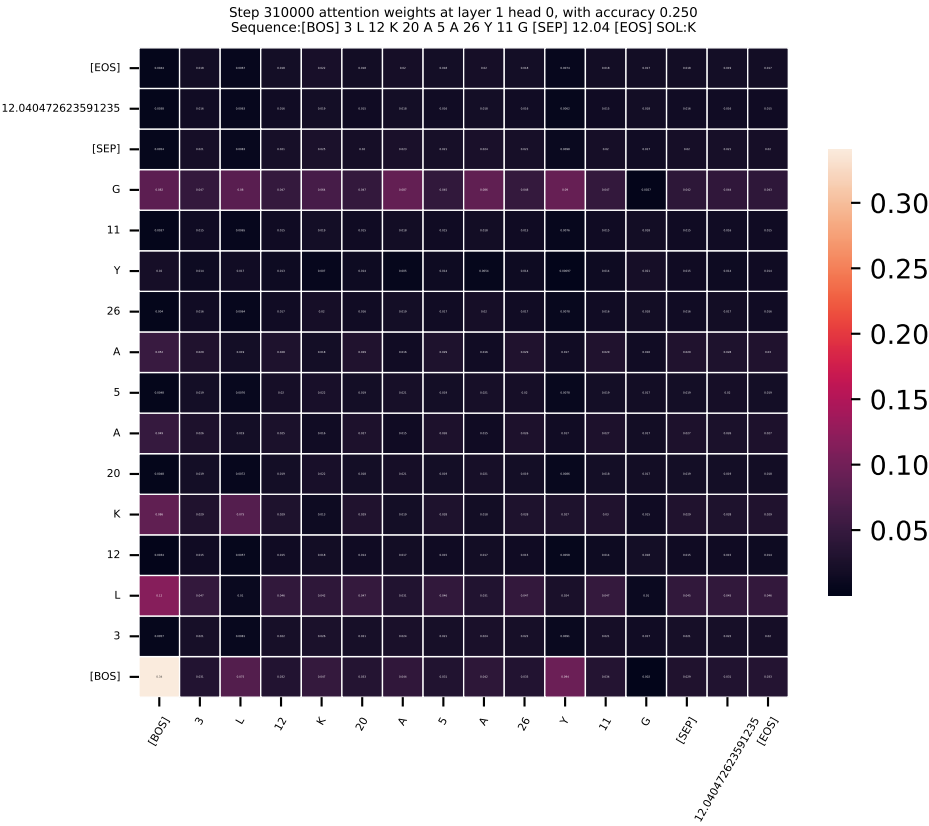


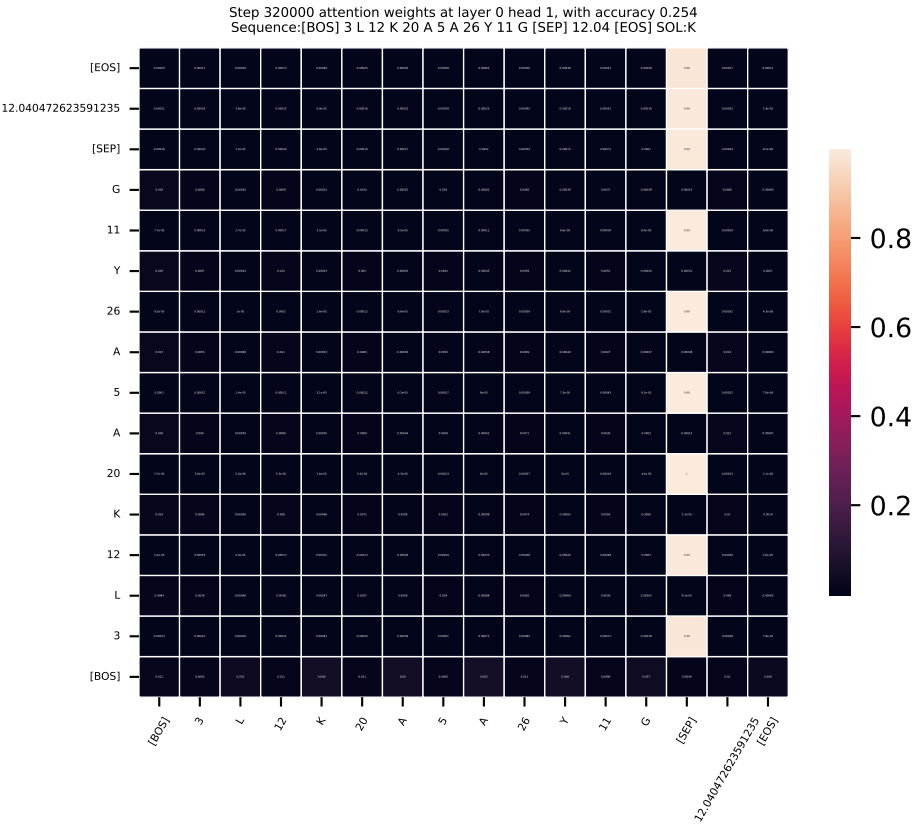
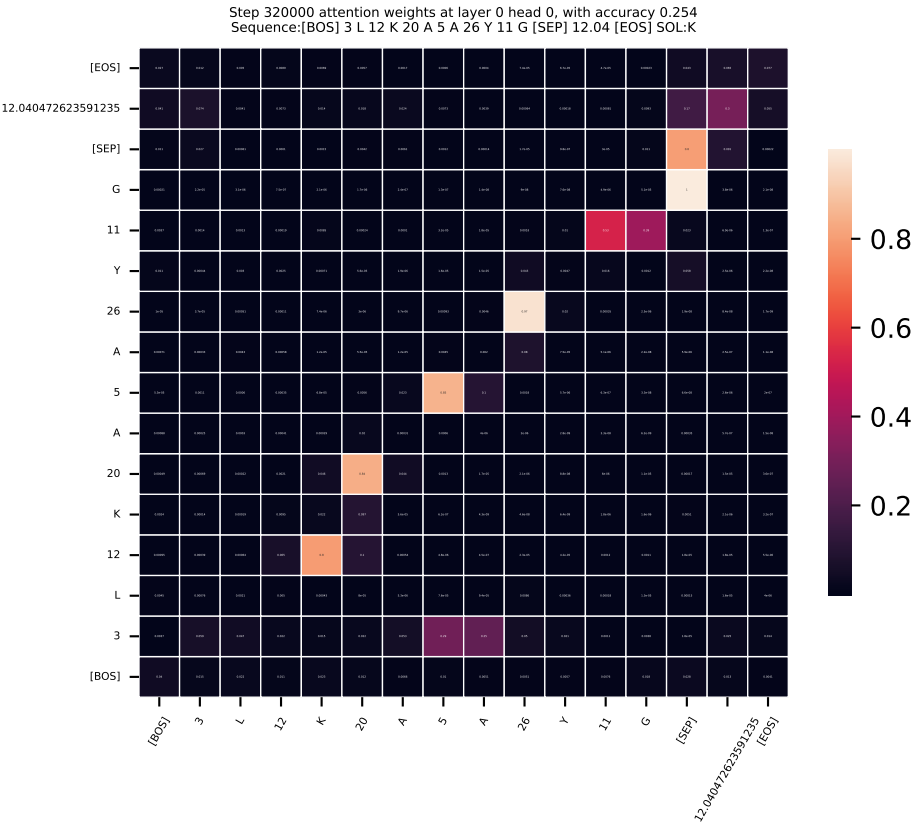


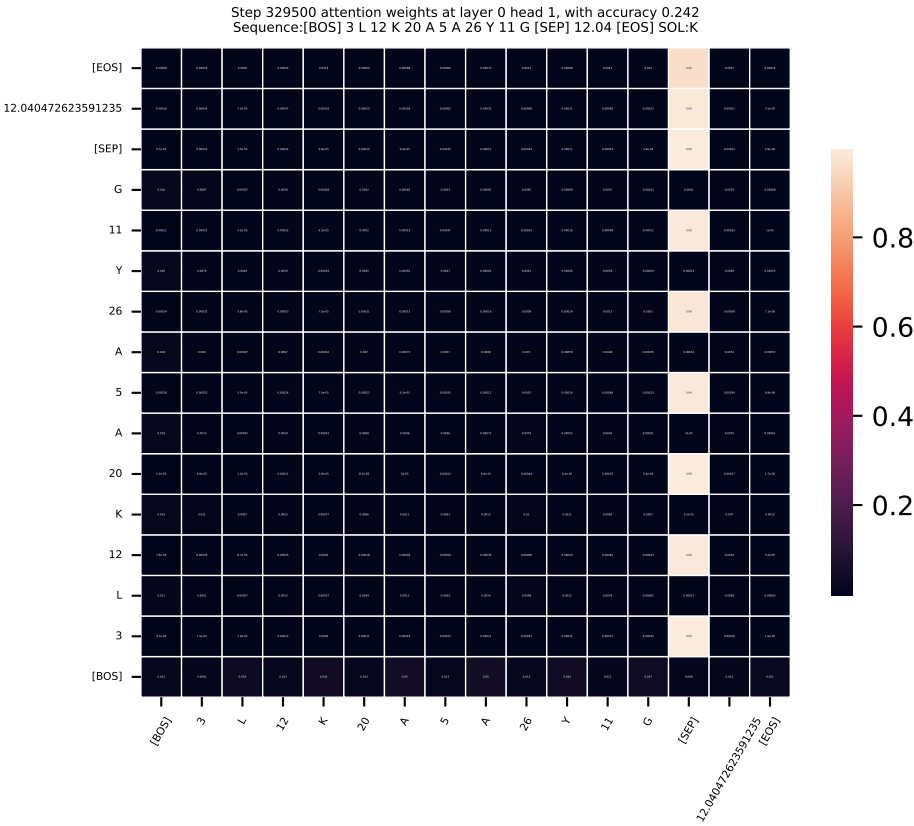
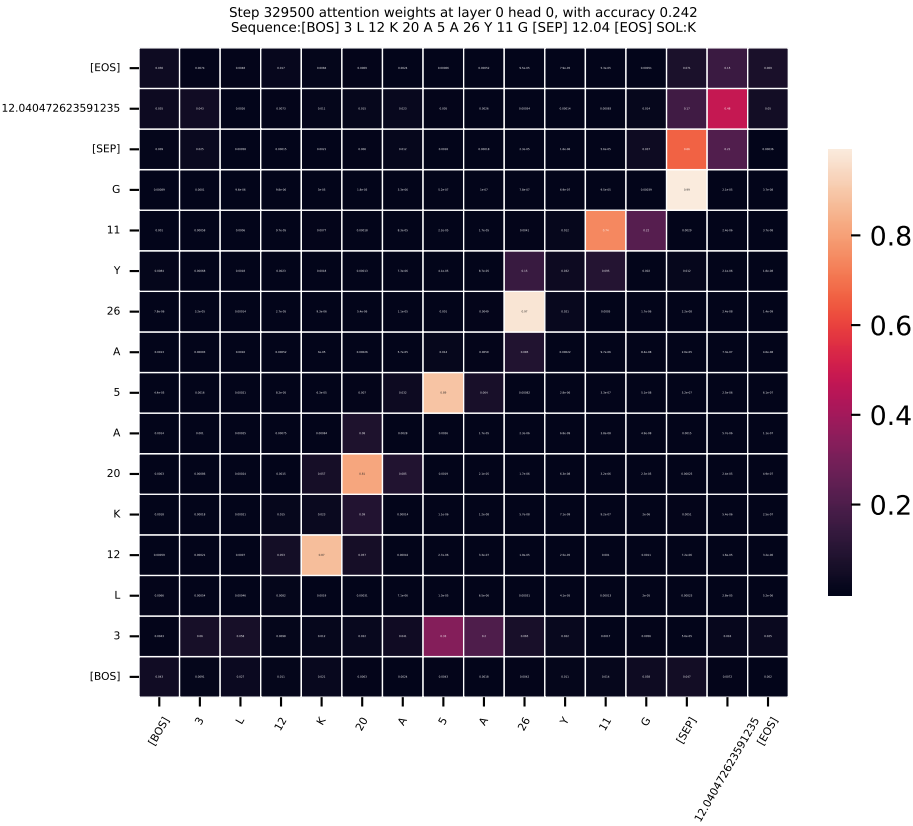


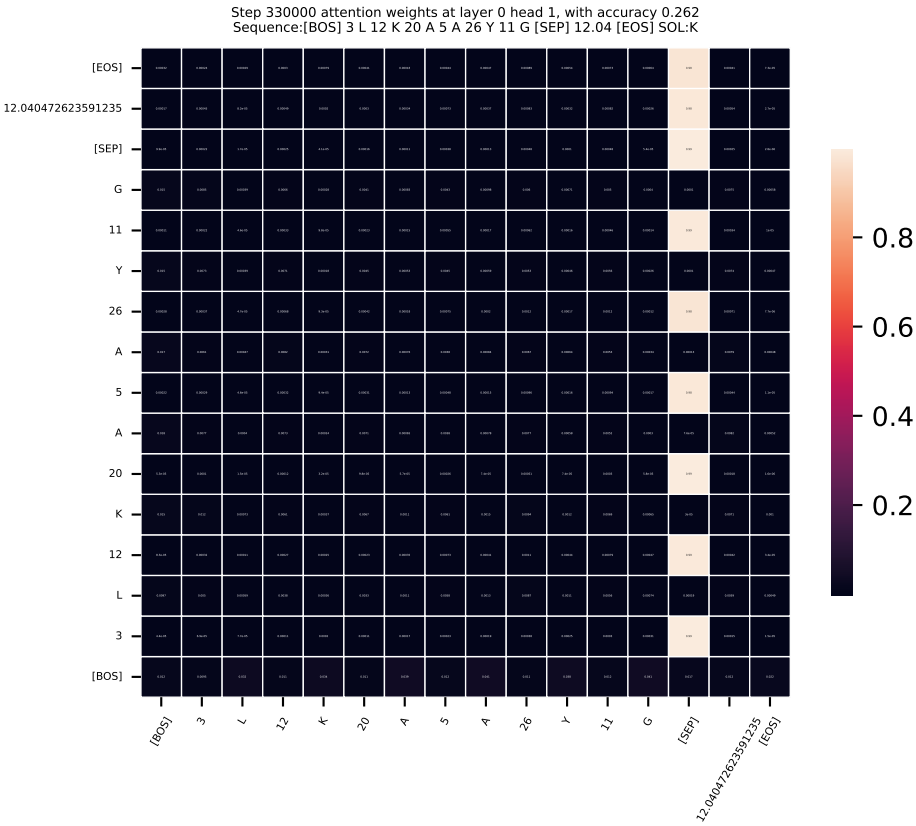
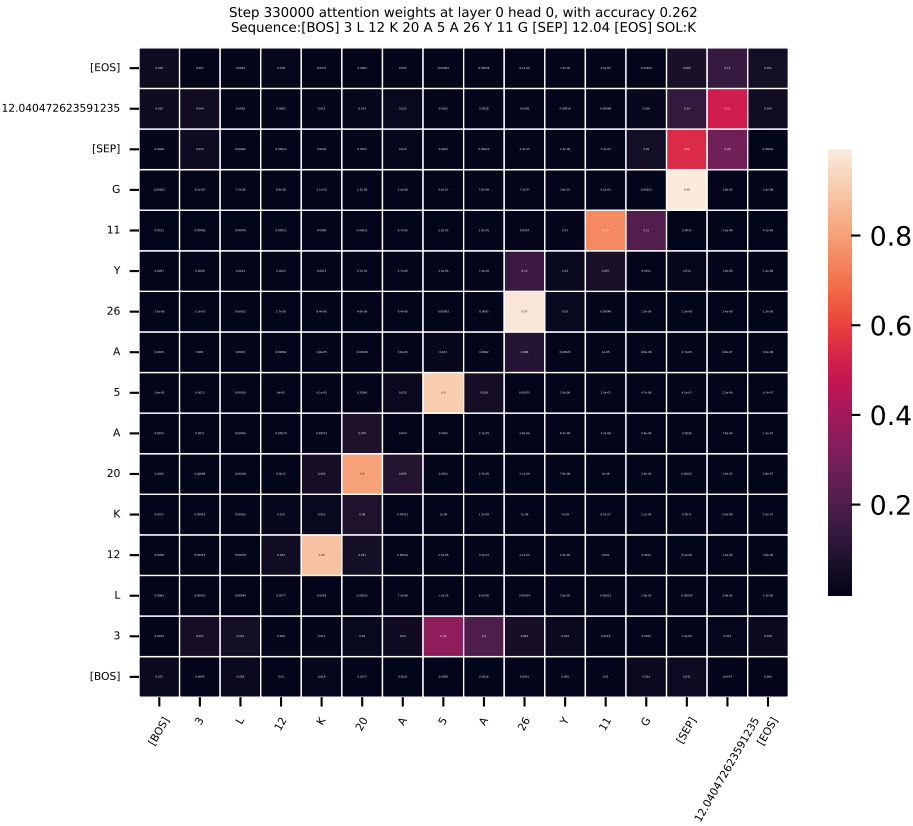


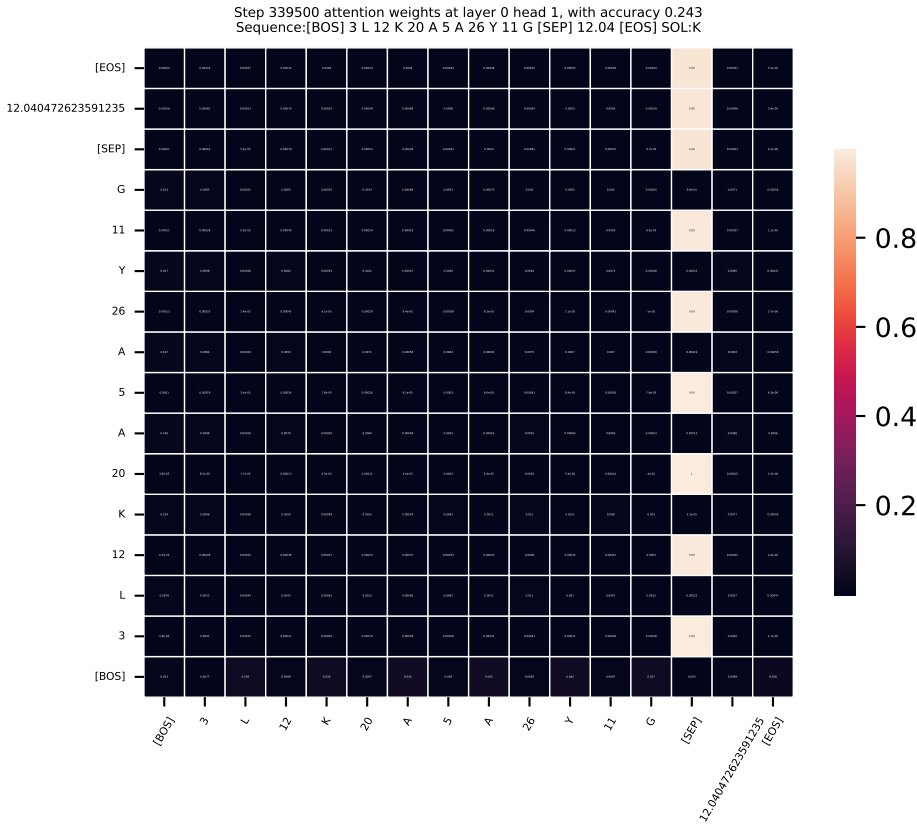
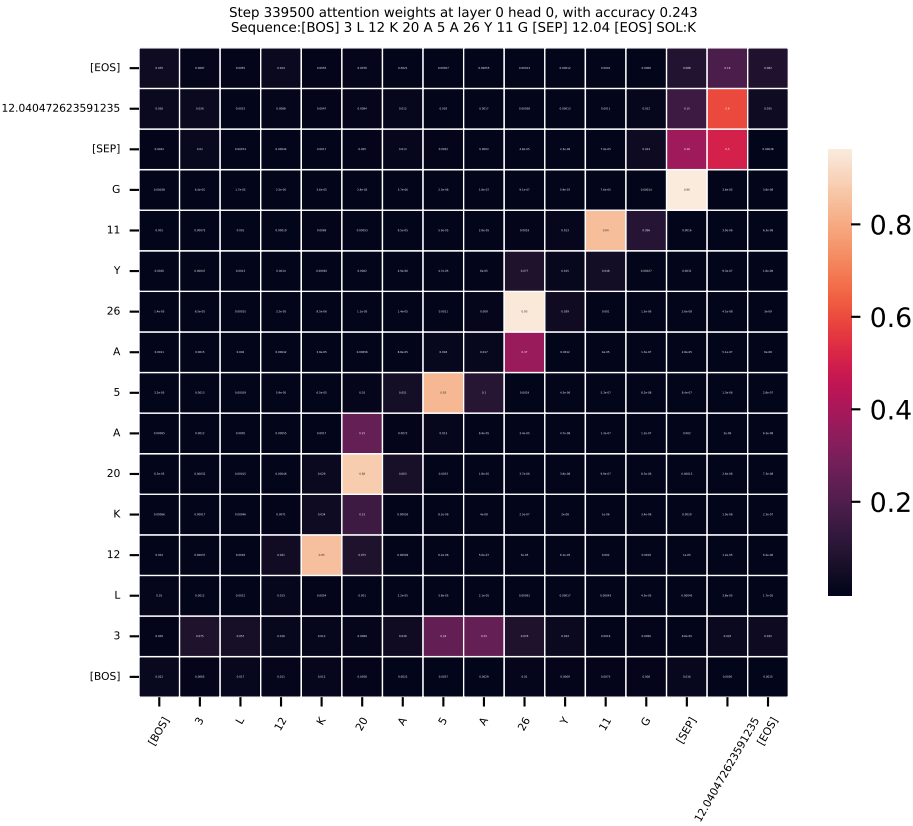


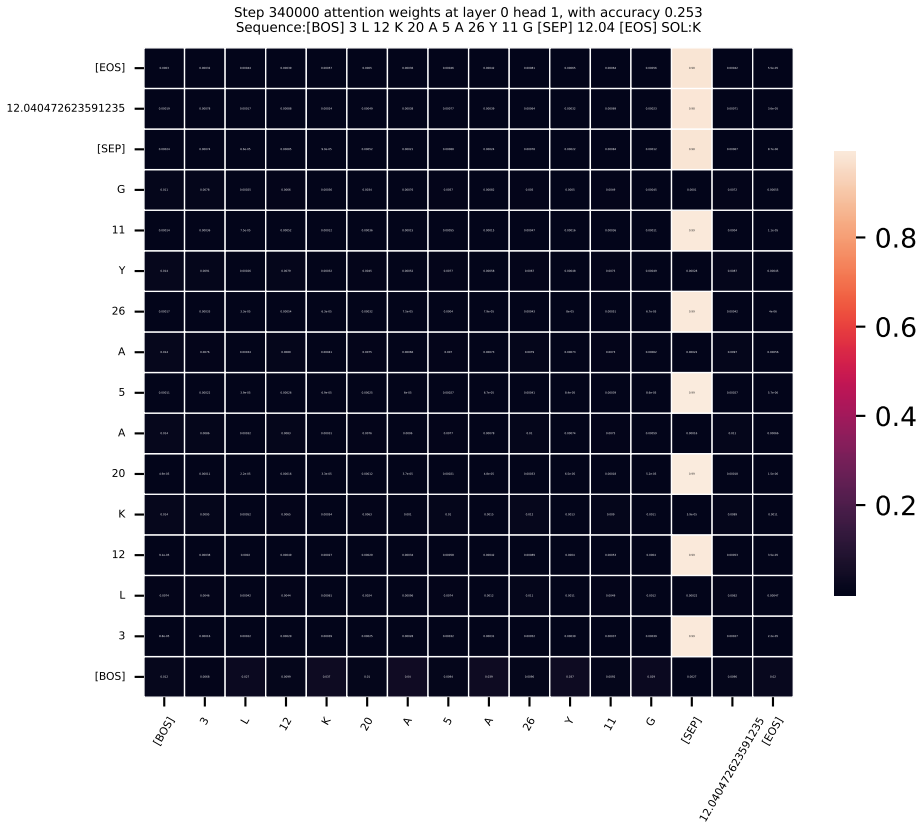
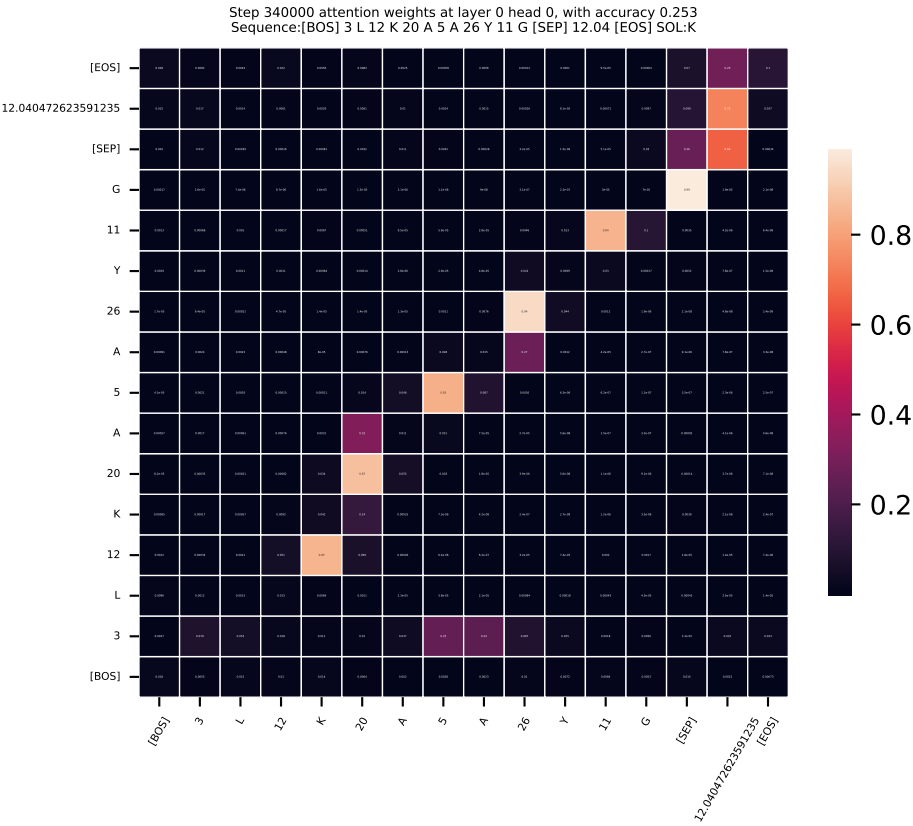


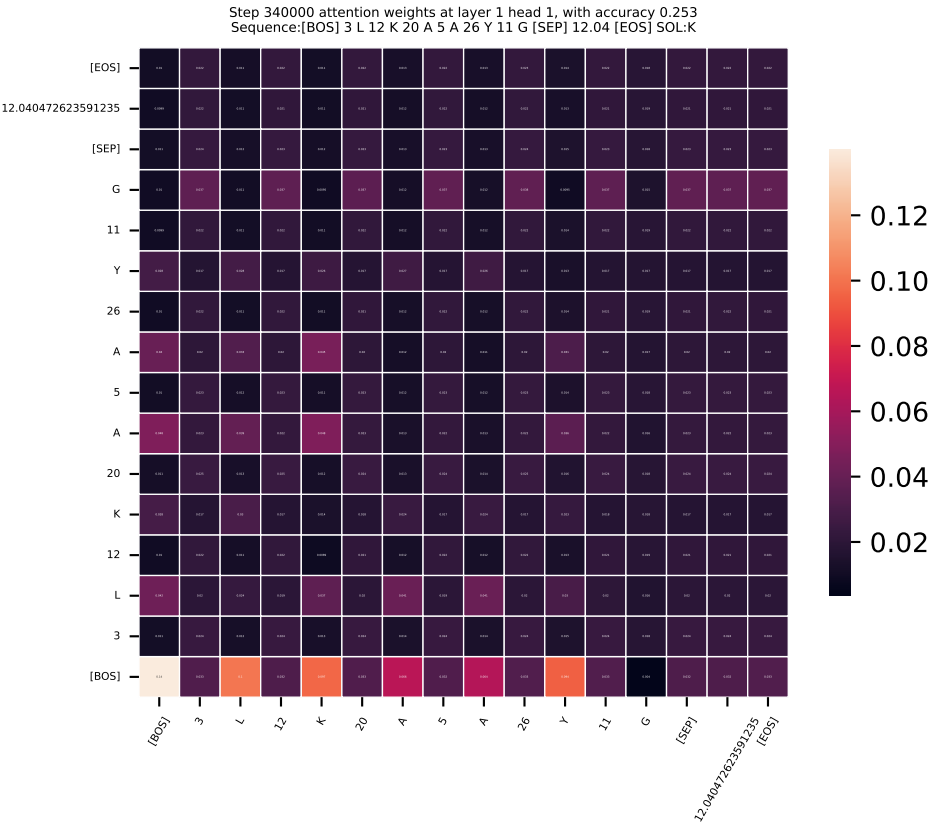
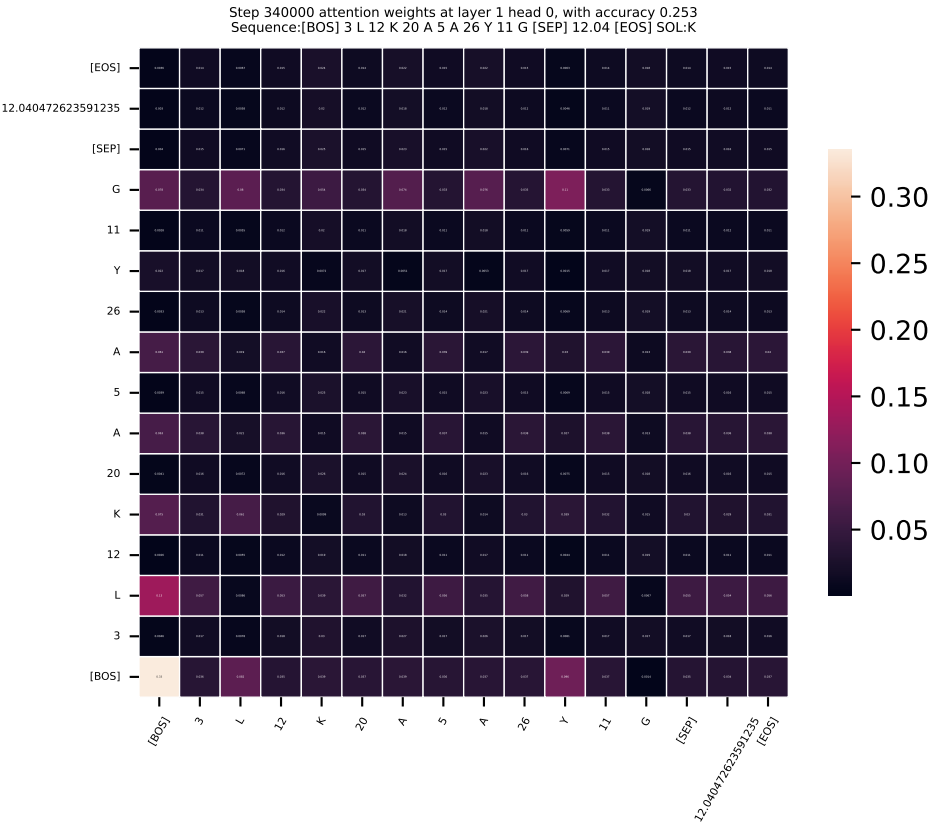




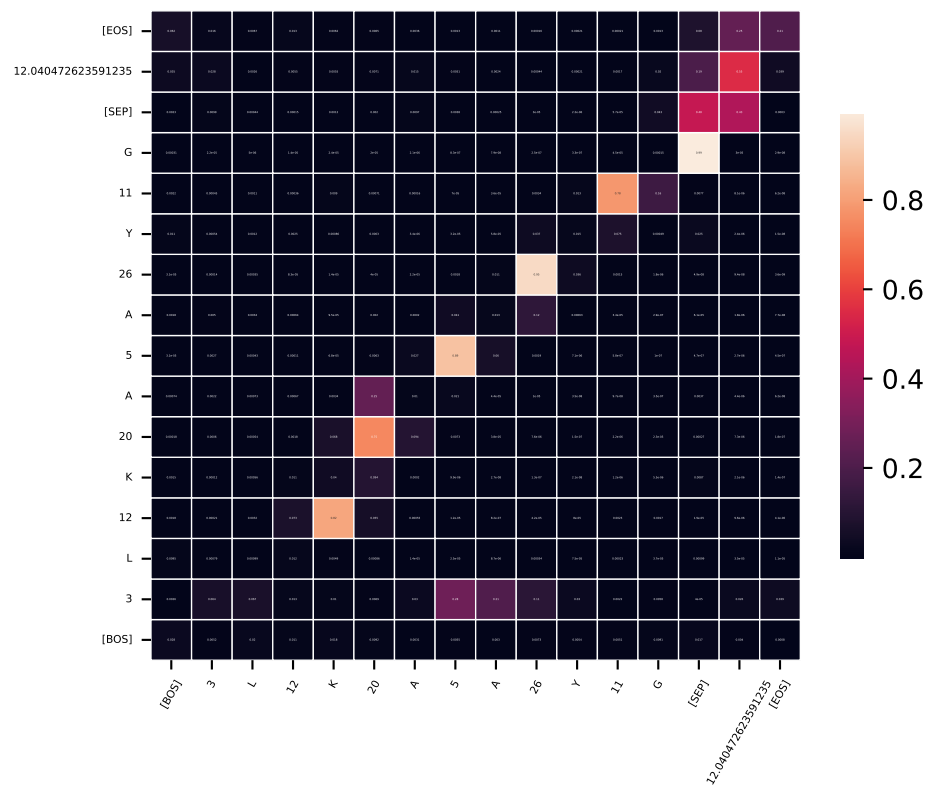








Step 349500 attention weights at layer 0 head 0, with accuracy 0.244
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K



Step 349500 attention weights at layer 0 head 1, with accuracy 0.244
Sequence:[BOS] 3 L 12 K 20 A 5 A 26 Y 11 G [SEP] 12.04 [EOS] SOL:K

