One-hot encoder with arch 399->800->Relu->200->Relu->32 on test set

Reconstruction error during training

Aminoacid changes (5 most common)
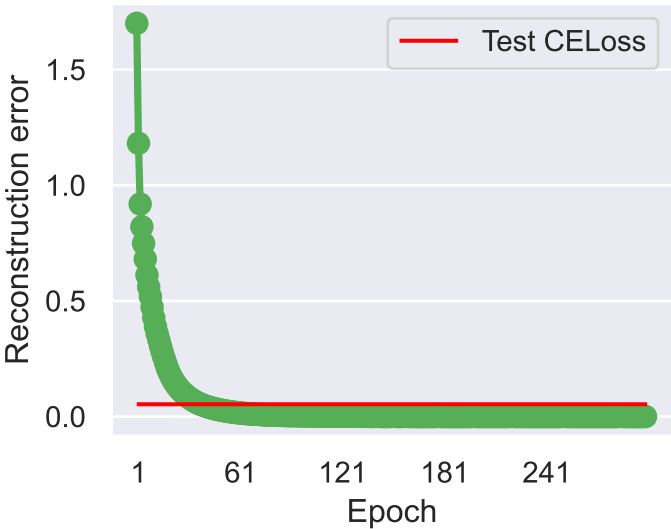
Average reconstruction error on test set:
CELoss: 0.05303

From 7057 sequences with right seq len:7038
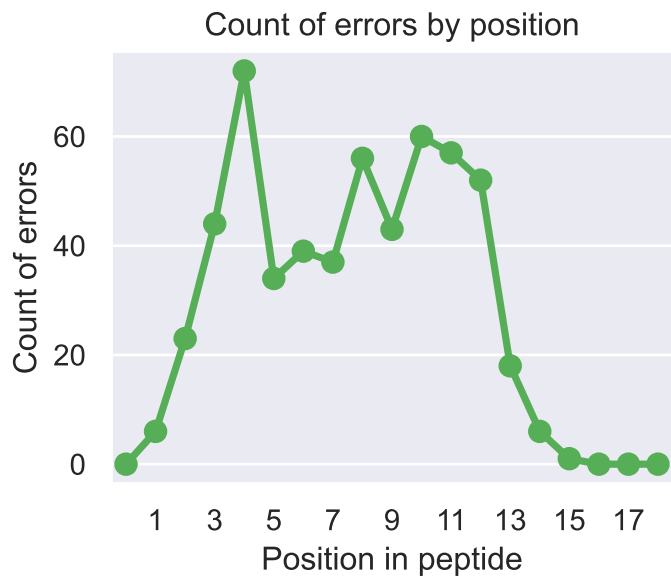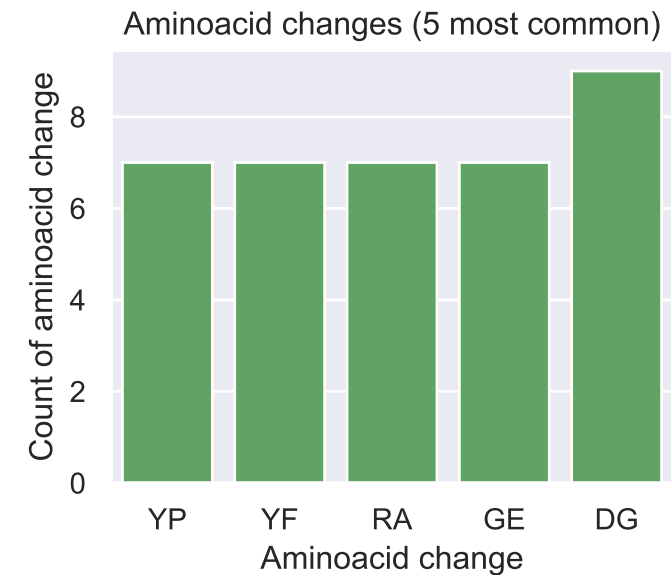
	precision	recall	f1-score	support
A	1.00	1.00	1.00	10527
C	1.00	1.00	1.00	7074
D	0.99	0.99	0.99	2845
E	0.99	1.00	1.00	5996
F	1.00	1.00	1.00	10894
G	1.00	1.00	1.00	10247
H	0.98	0.97	0.98	1070
I	0.98	0.97	0.97	1229
K	0.96	0.96	0.96	679
L	0.99	0.99	0.99	4367
M	0.95	0.94	0.95	352
N	0.99	0.99	0.99	3016
P	0.99	0.99	0.99	2993
Q	1.00	1.00	1.00	6408
R	0.99	0.99	0.99	3162
S	1.00	1.00	1.00	15847
T	0.99	1.00	1.00	6482
V	0.99	0.99	0.99	1909
W	0.98	0.96	0.97	655
Y	0.99	0.99	0.99	6125
accuracy			0.99	101877
macro avg	0.99	0.99	0.99	101877
weighted avg	0.99	0.99	0.99	101877

Count of errors by position

Count of errors by position (in %)