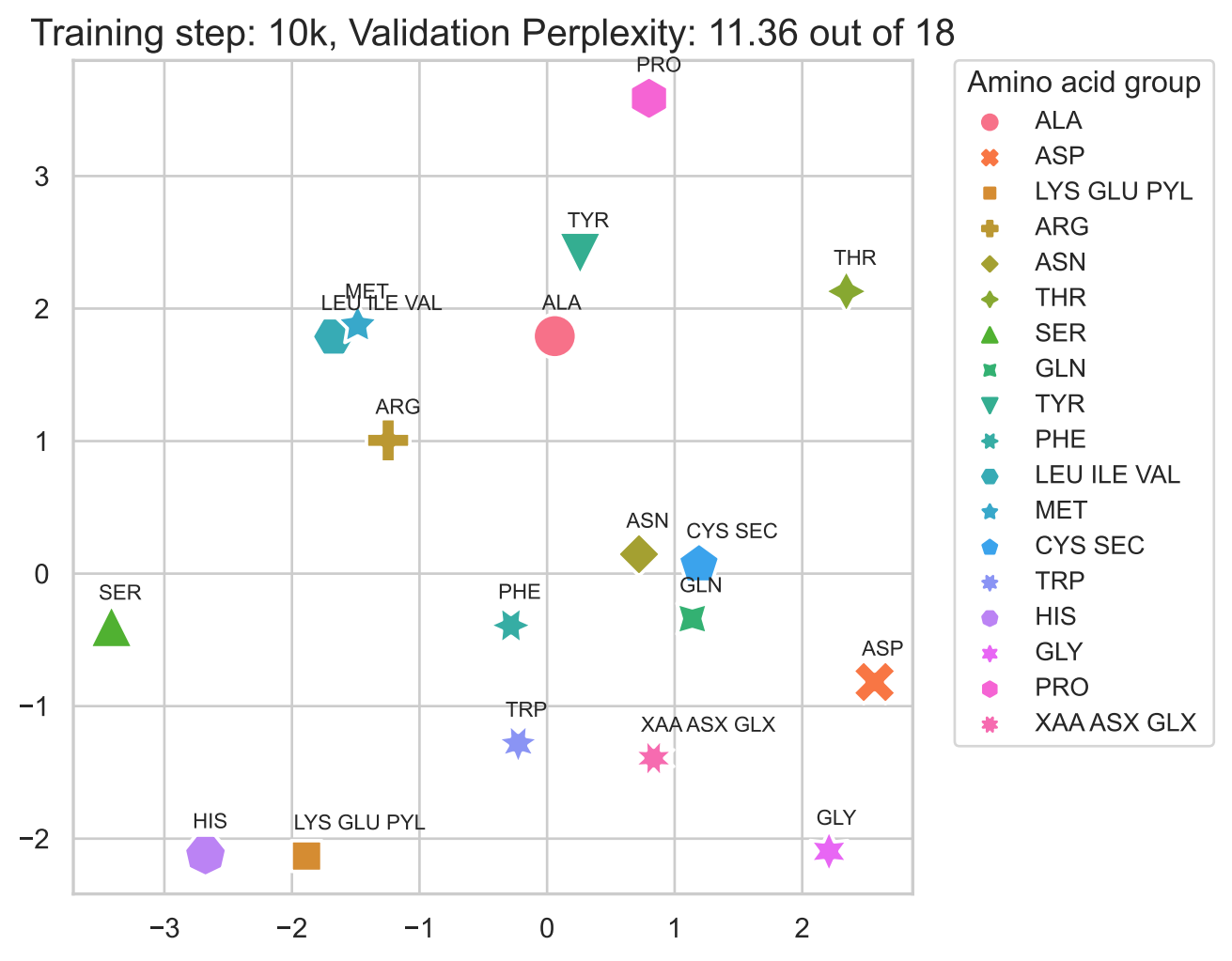
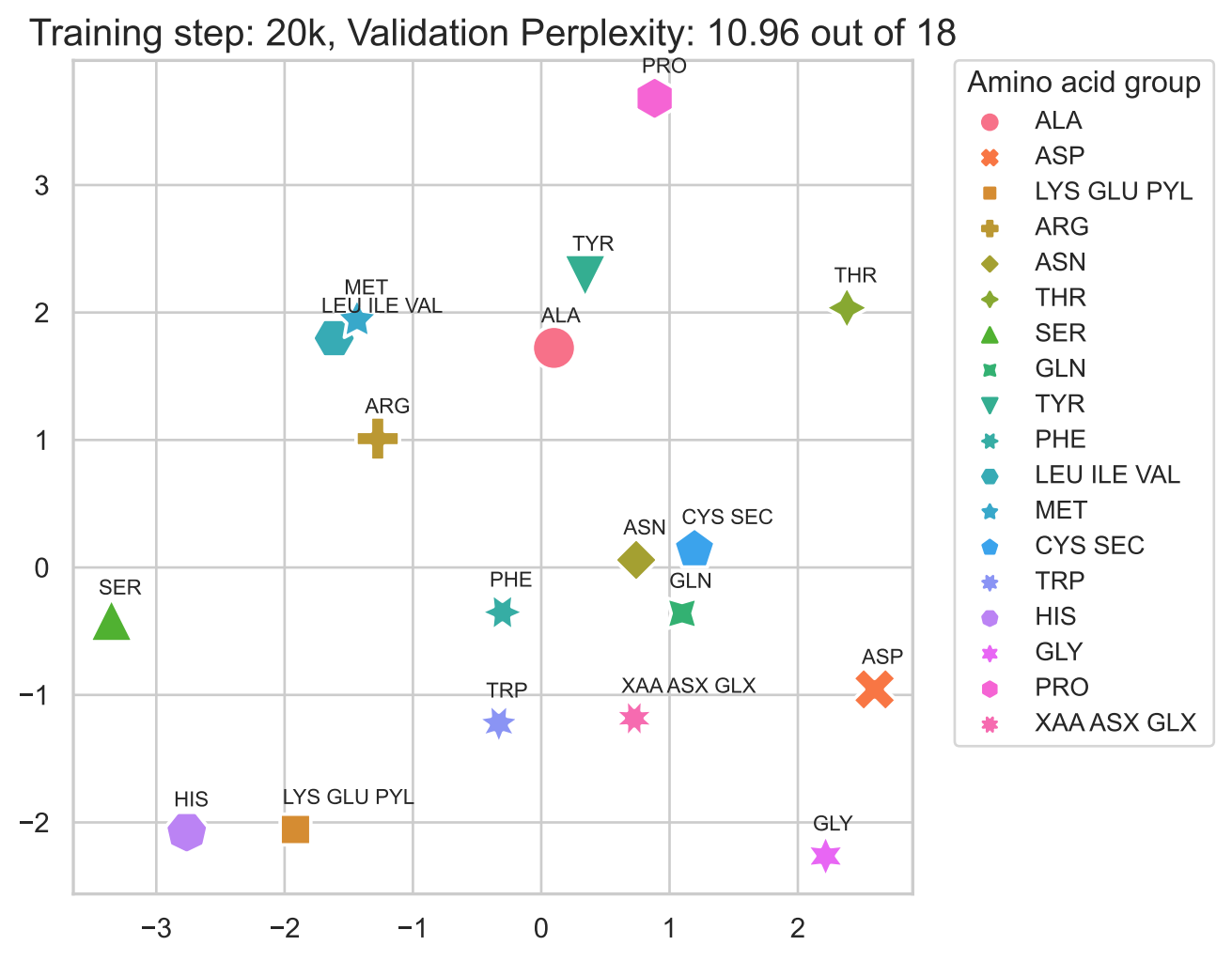
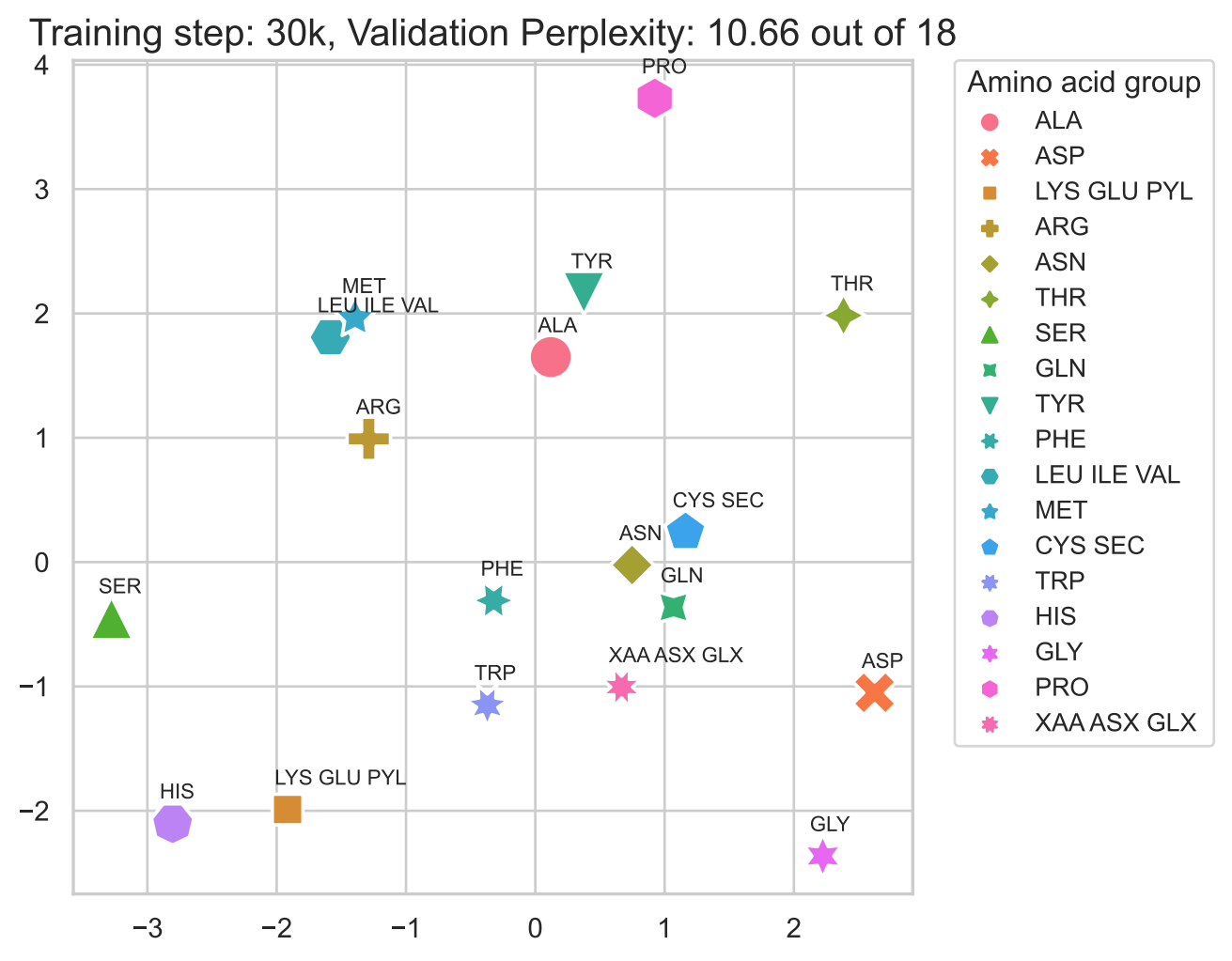
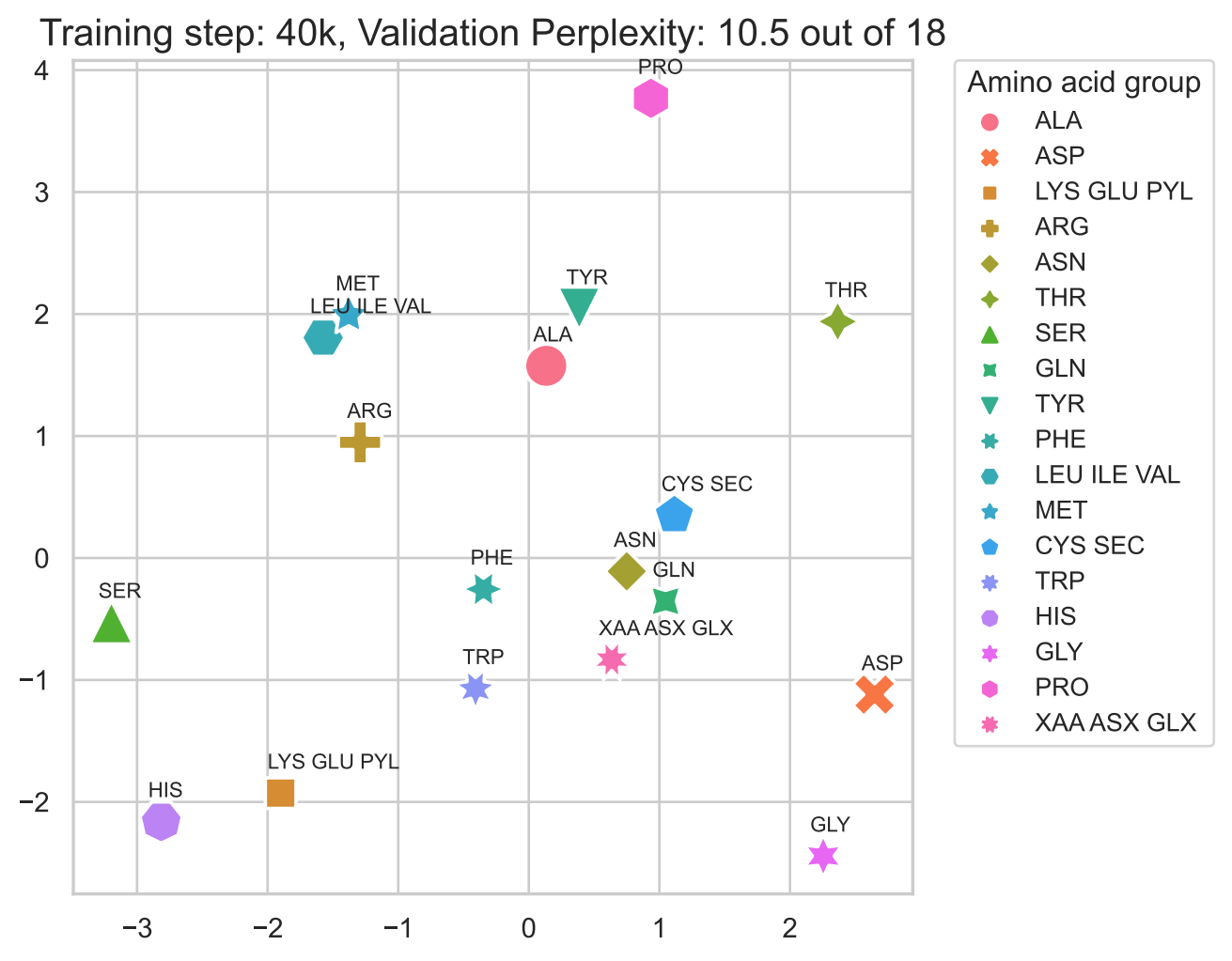
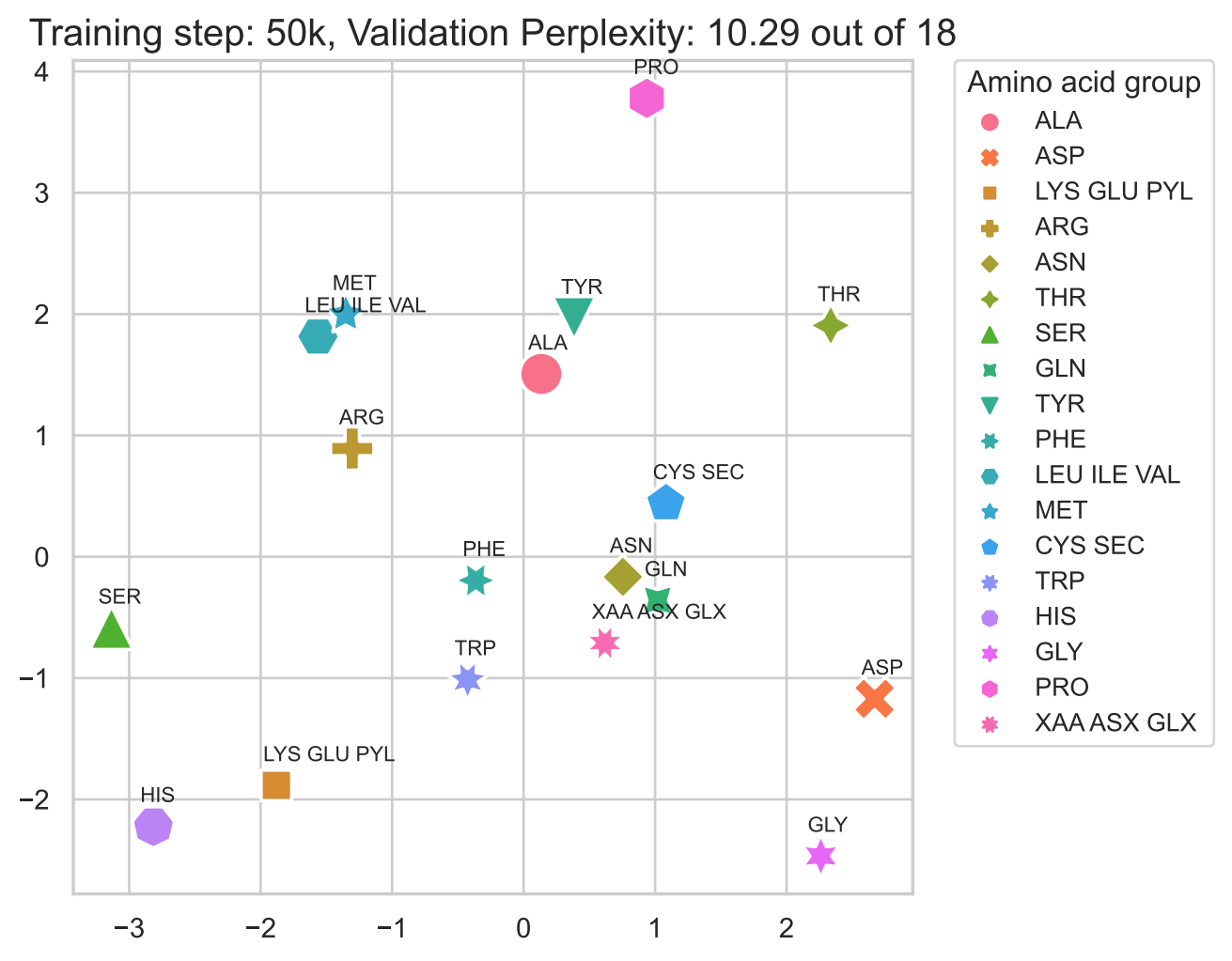
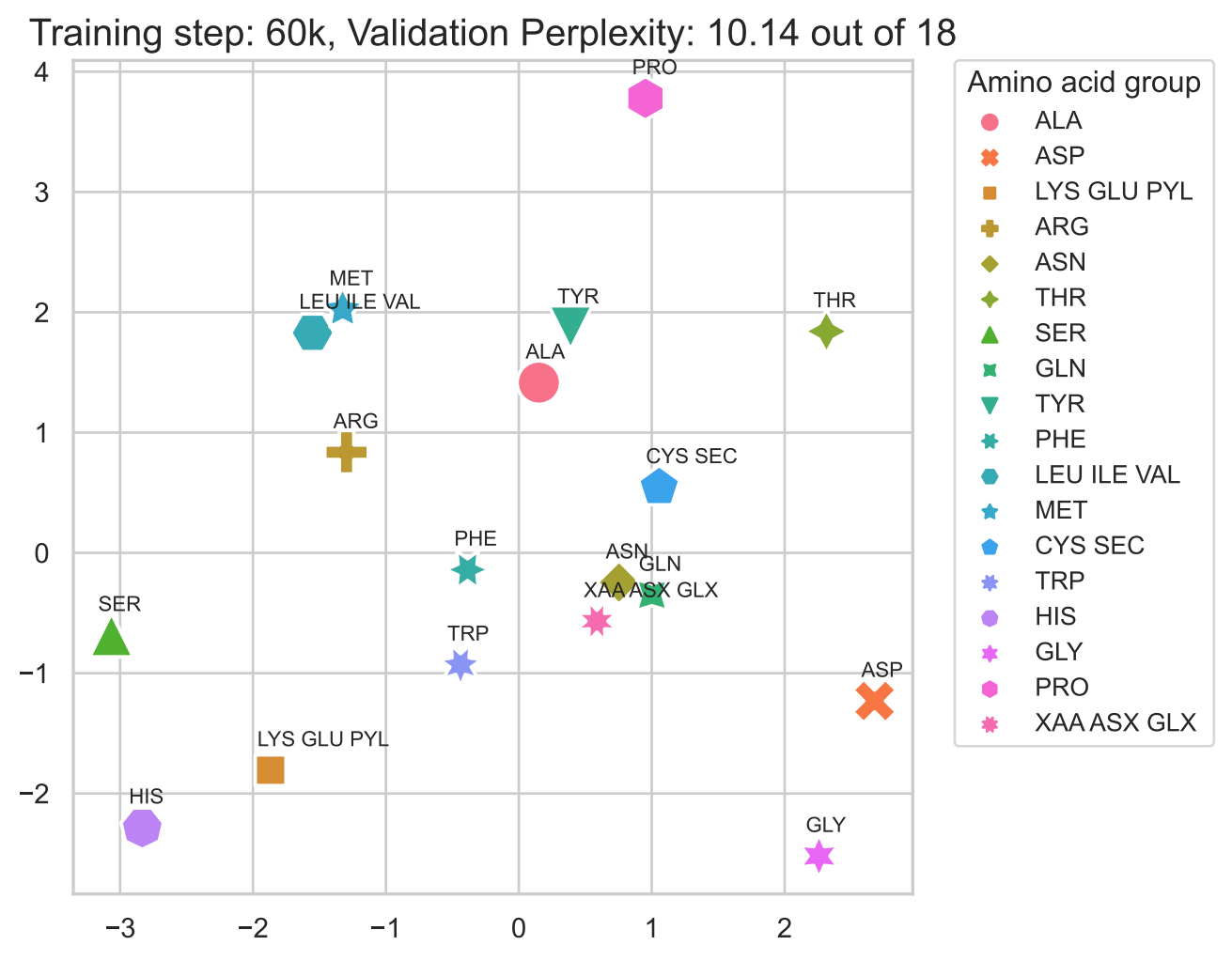

Training step: 10k, Validation Perplexity: 11.36 out of 18
PRO
TYR
THR
MET
LEU ILE VAL
ALA
ARG
ASN
CYS SEC
SER
PHE
GLN
ASP
TRP
XAA ASX GLX
HIS
LYS GLU PYL
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

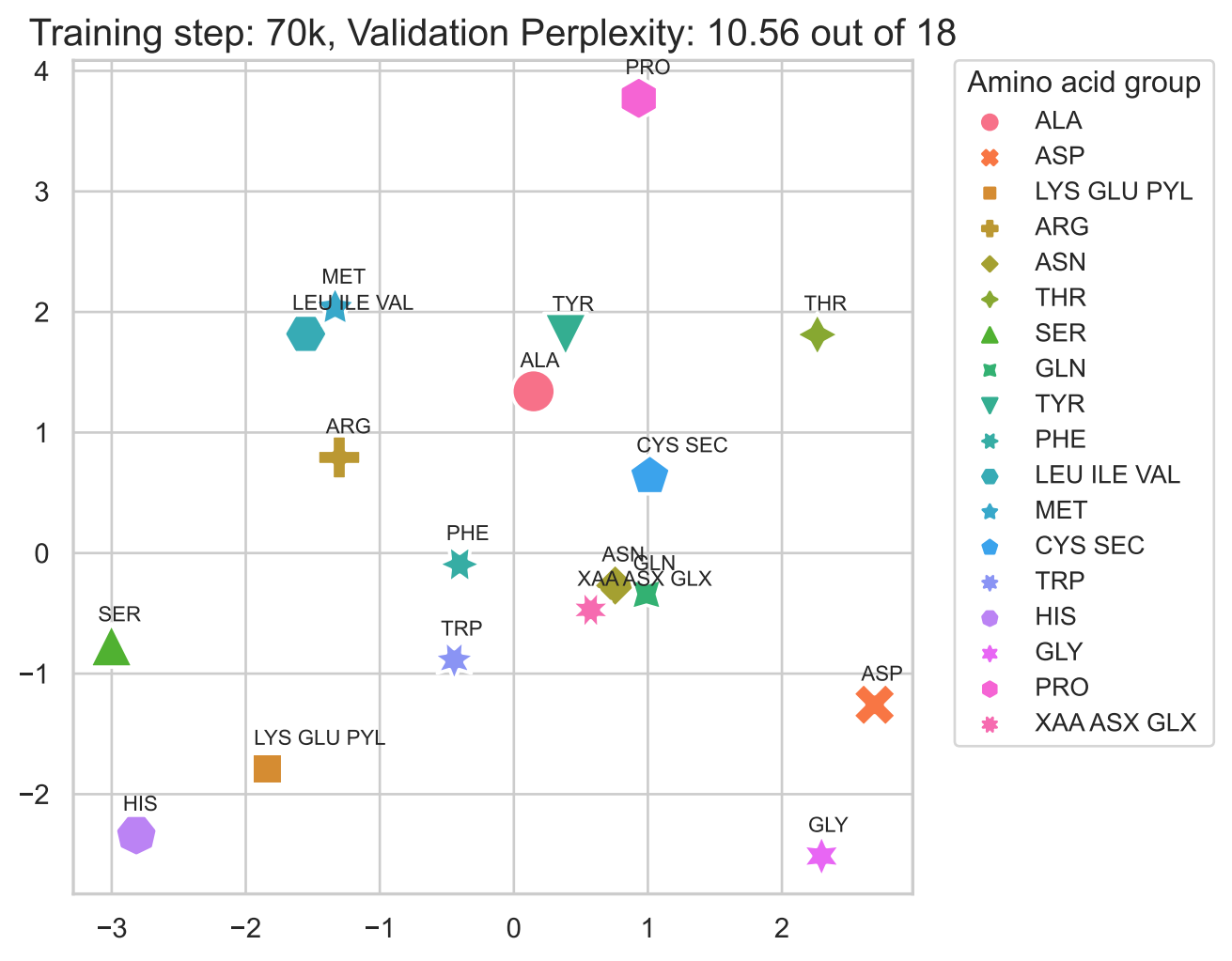

Training step: 20k, Validation Perplexity: 10.96 out of 18
PRO
TYR
THR
MET
LEU ILE VAL
ALA
ARG
ASN
CYS SEC
SER
PHE
GLN
ASP
TRP
XAA ASX GLX
HIS
LYS GLU PYL
GLY
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

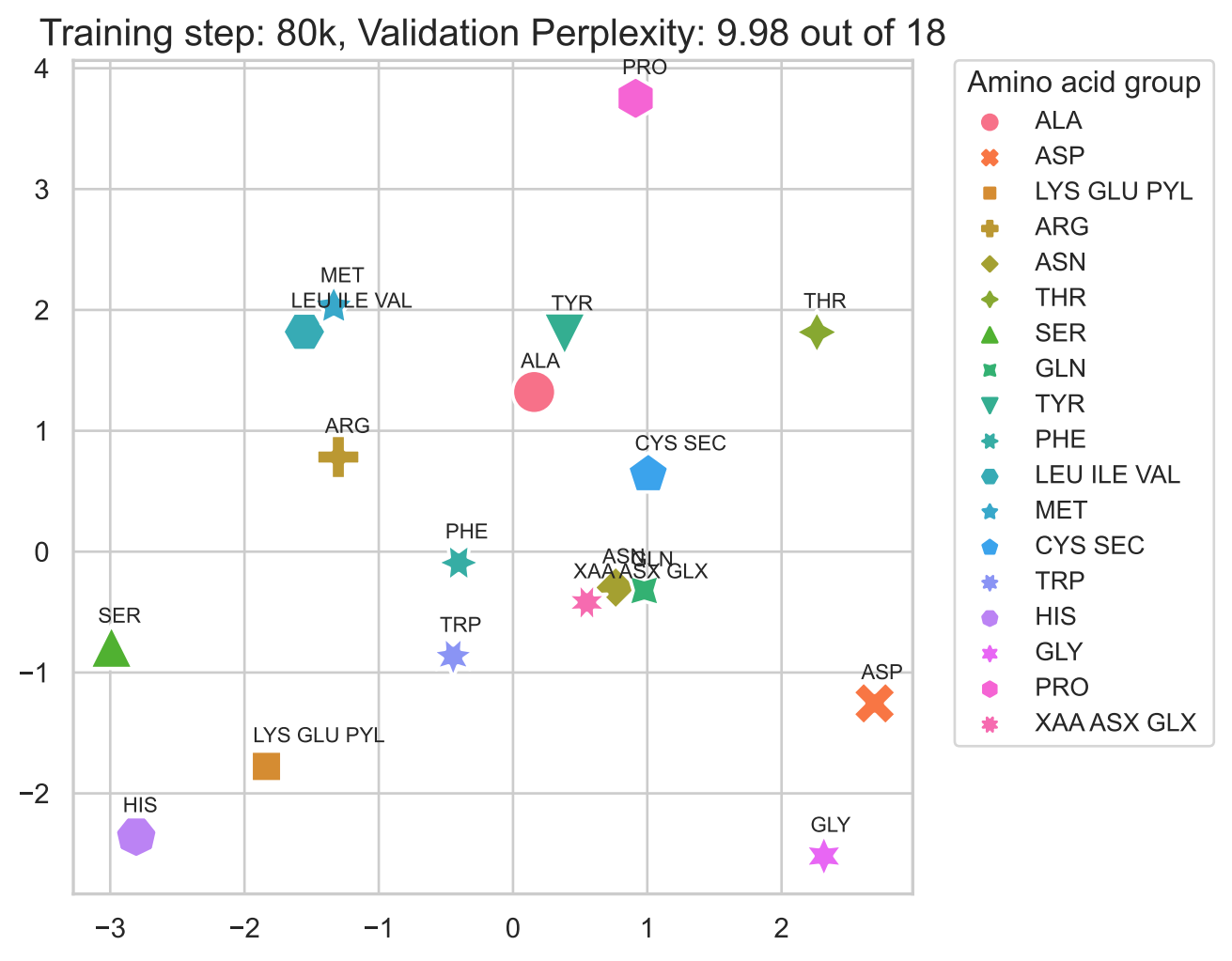

Training step: 30k, Validation Perplexity: 10.66 out of 18
PRO
TYR
MET
LEU ILE VAL
THR
ALA
ARG
CYS SEC
ASN
PHE
GLN
SER
XAA ASX GLX
ASP
TRP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

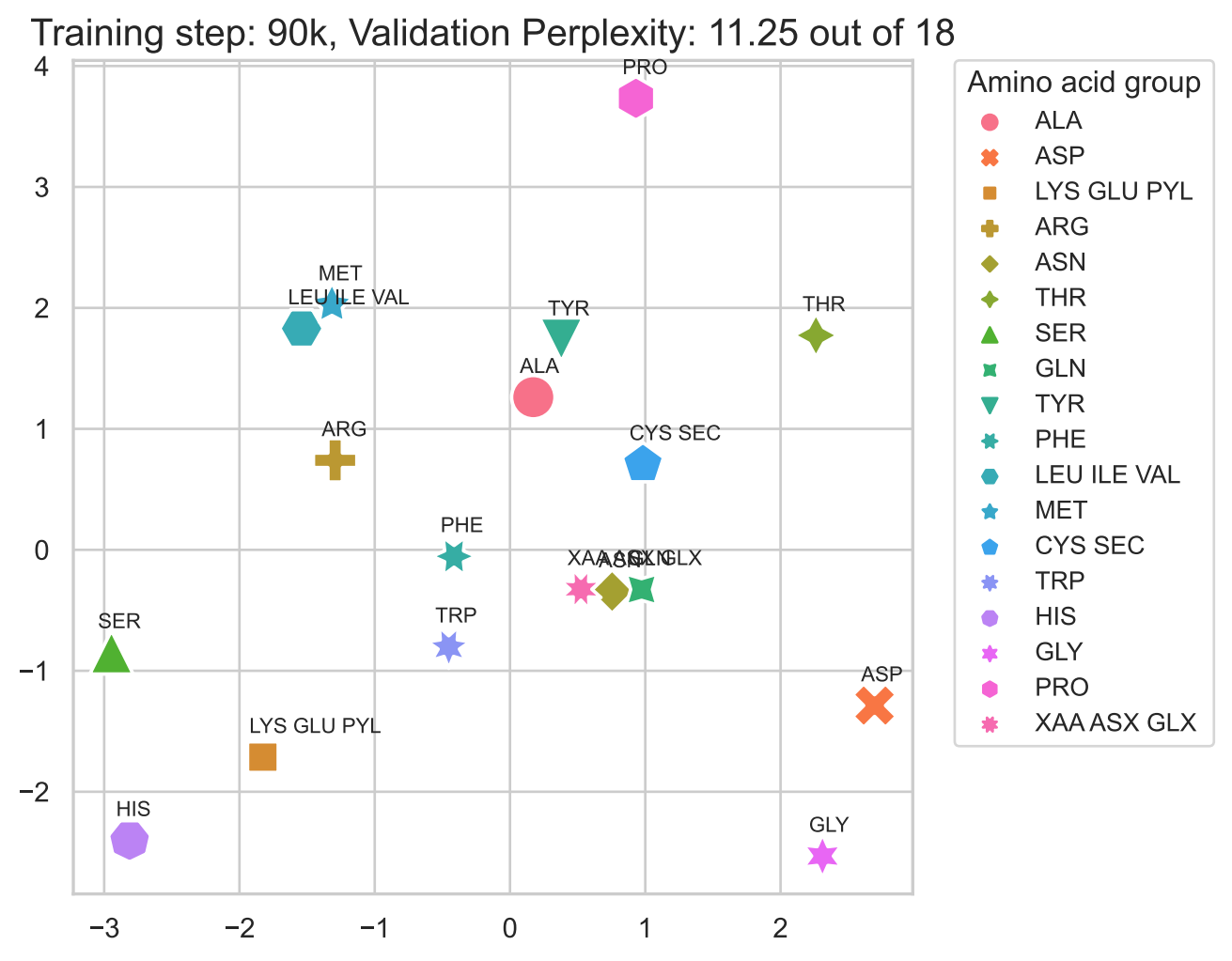

Training step: 40k, Validation Perplexity: 10.5 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
ASN
PHE
GLN
SER
XAA ASX GLX
TRP
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

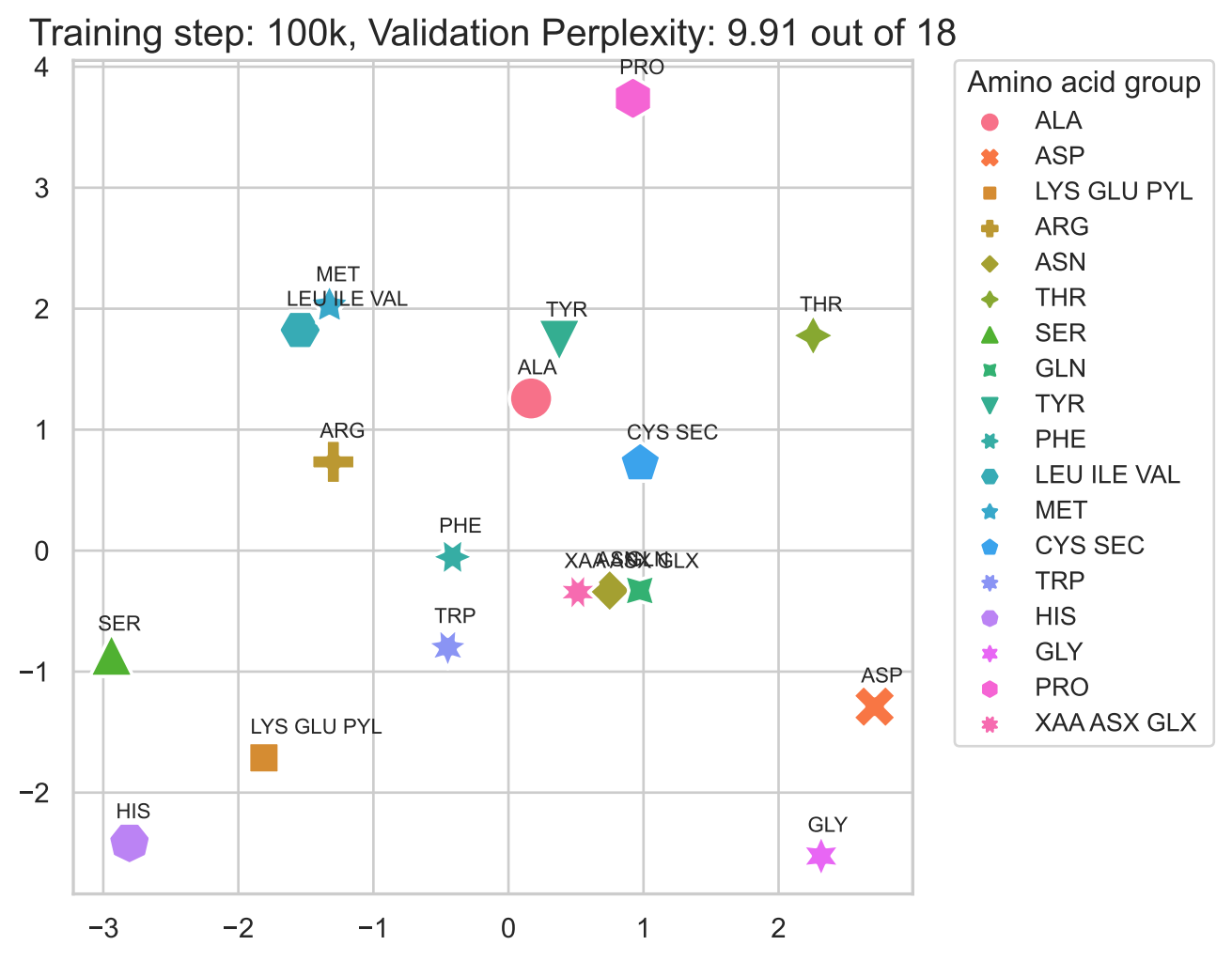

Training step: 50k, Validation Perplexity: 10.29 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
ASN
GLN
SER
XAA ASX GLX
TRP
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

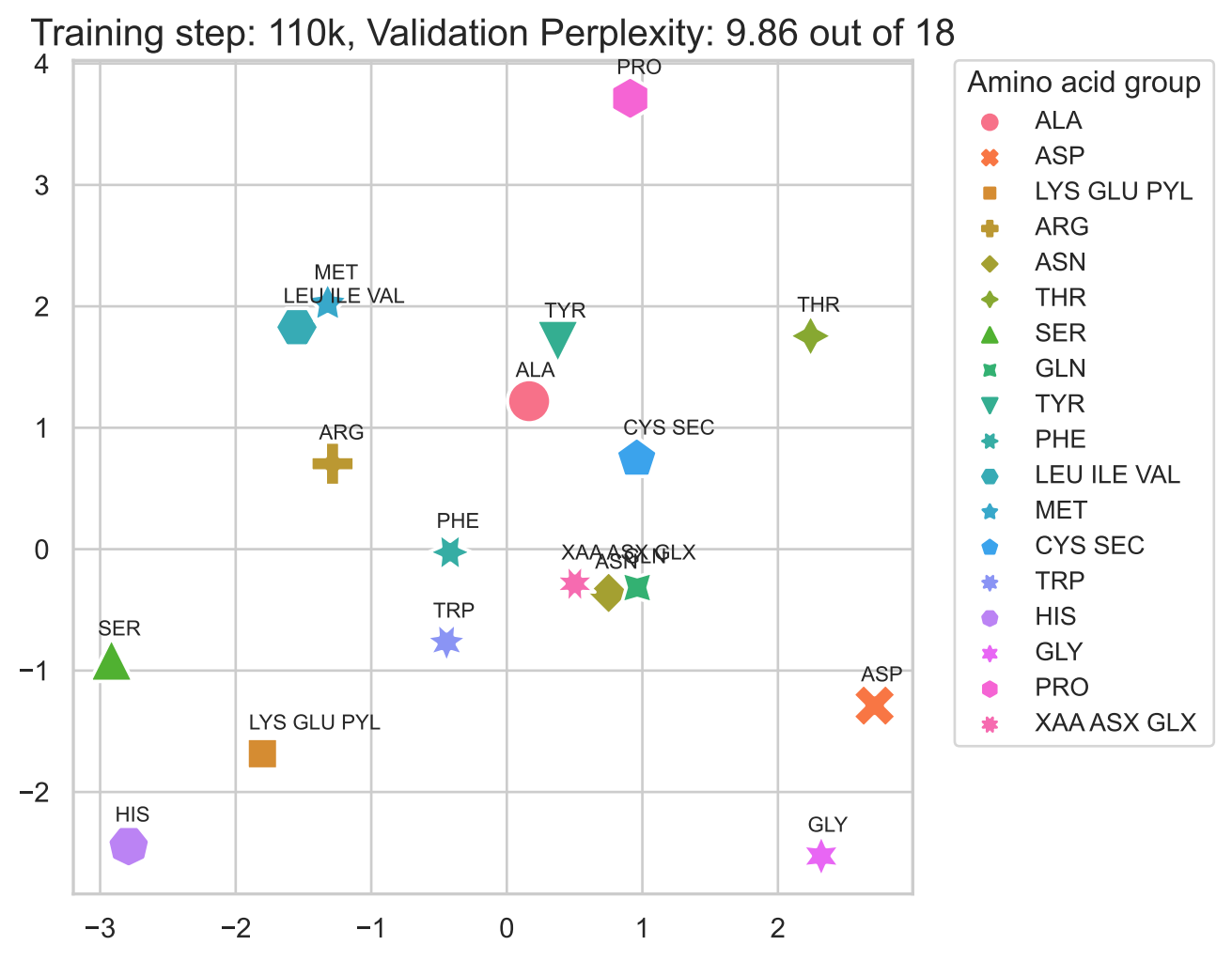

Training step: 60k, Validation Perplexity: 10.14 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
ASN
GLN
XAA ASX GLX
SER
TRP
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

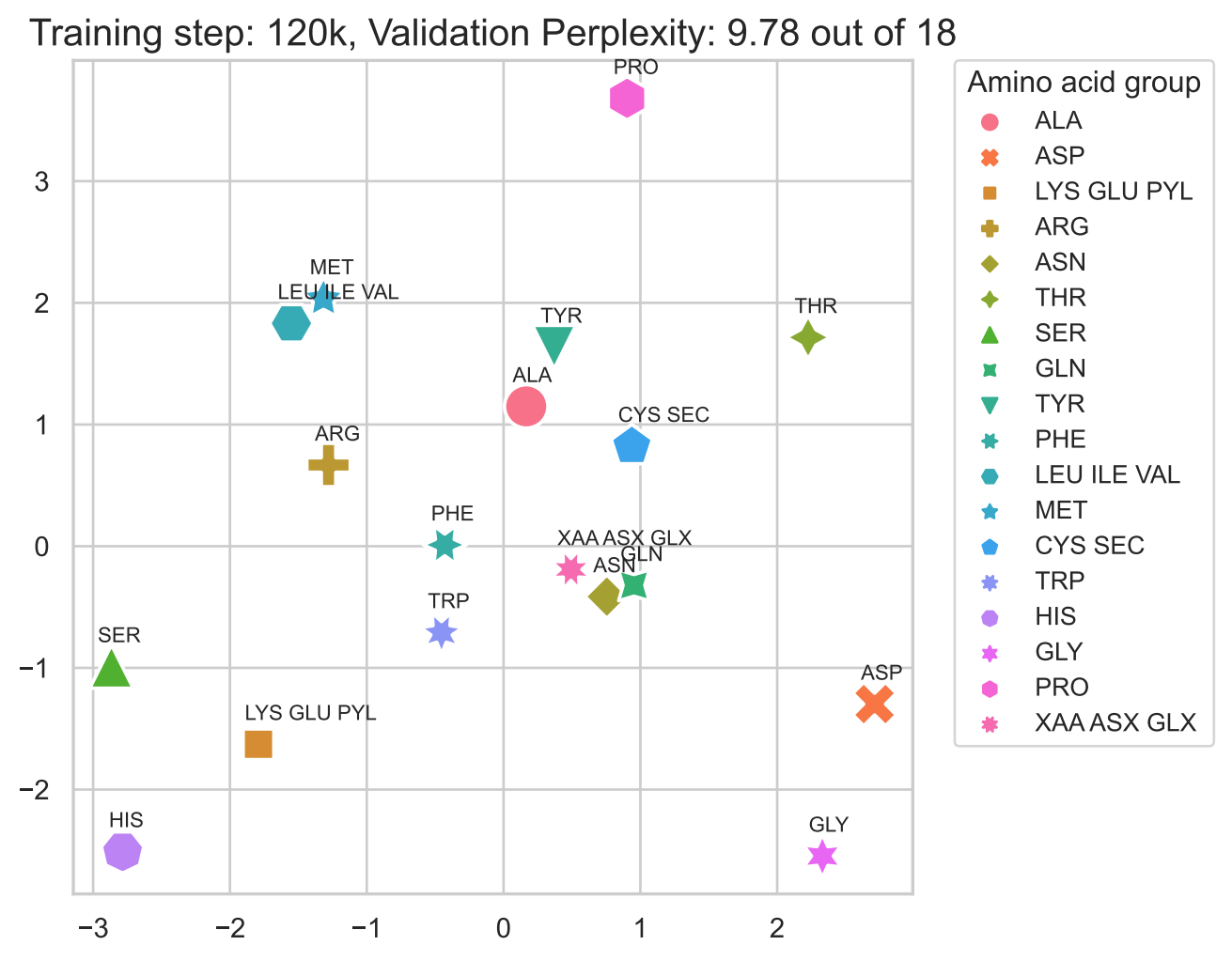

Training step: 70k, Validation Perplexity: 10.56 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
ASN
GLN
XAA ASX GLX
SER
TRP
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

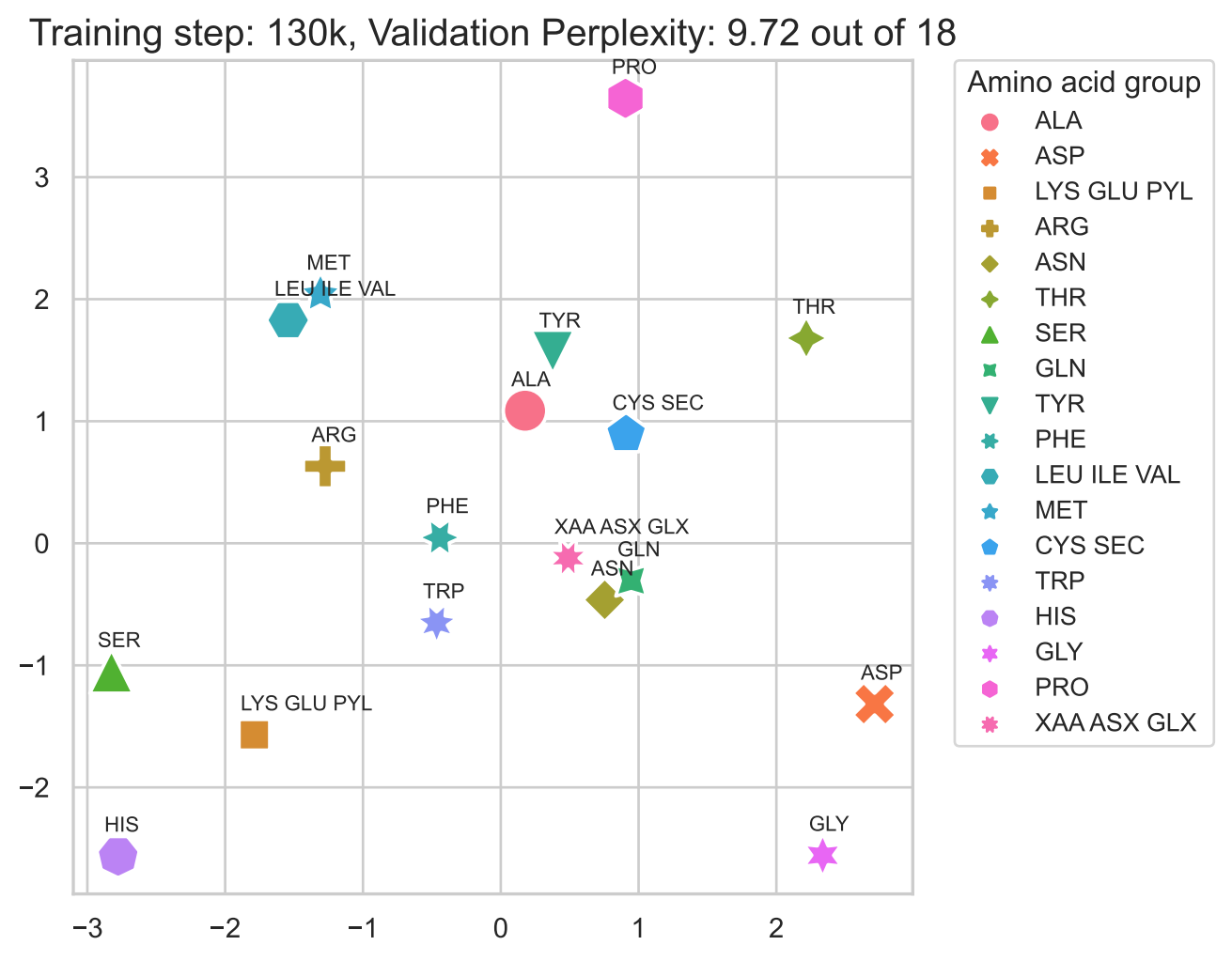

Training step: 80k, Validation Perplexity: 9.98 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
ASN
GLN
XAA ASX GLX
SER
TRP
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

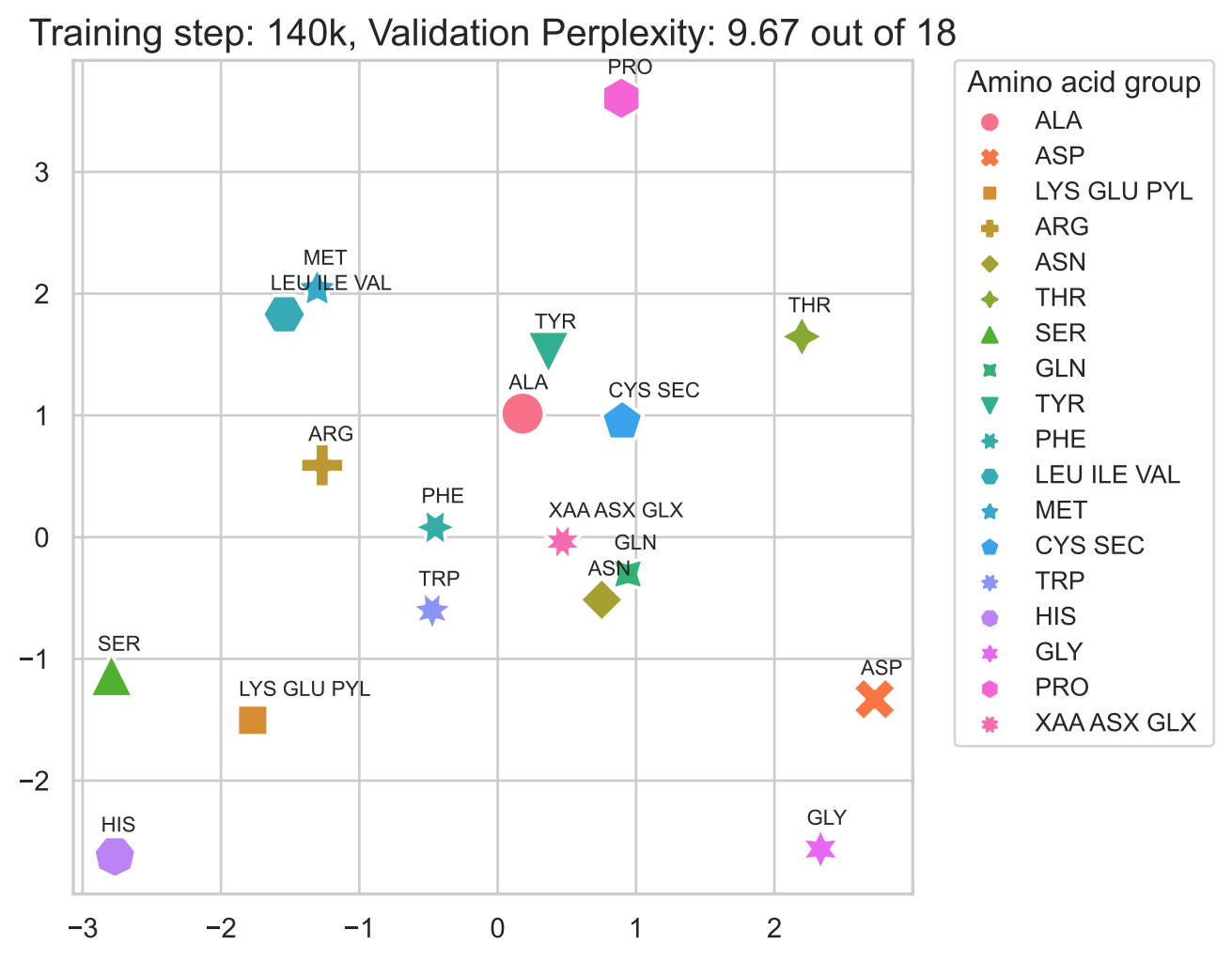

Training step: 90k, Validation Perplexity: 11.25 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
XAA ASX GLX
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

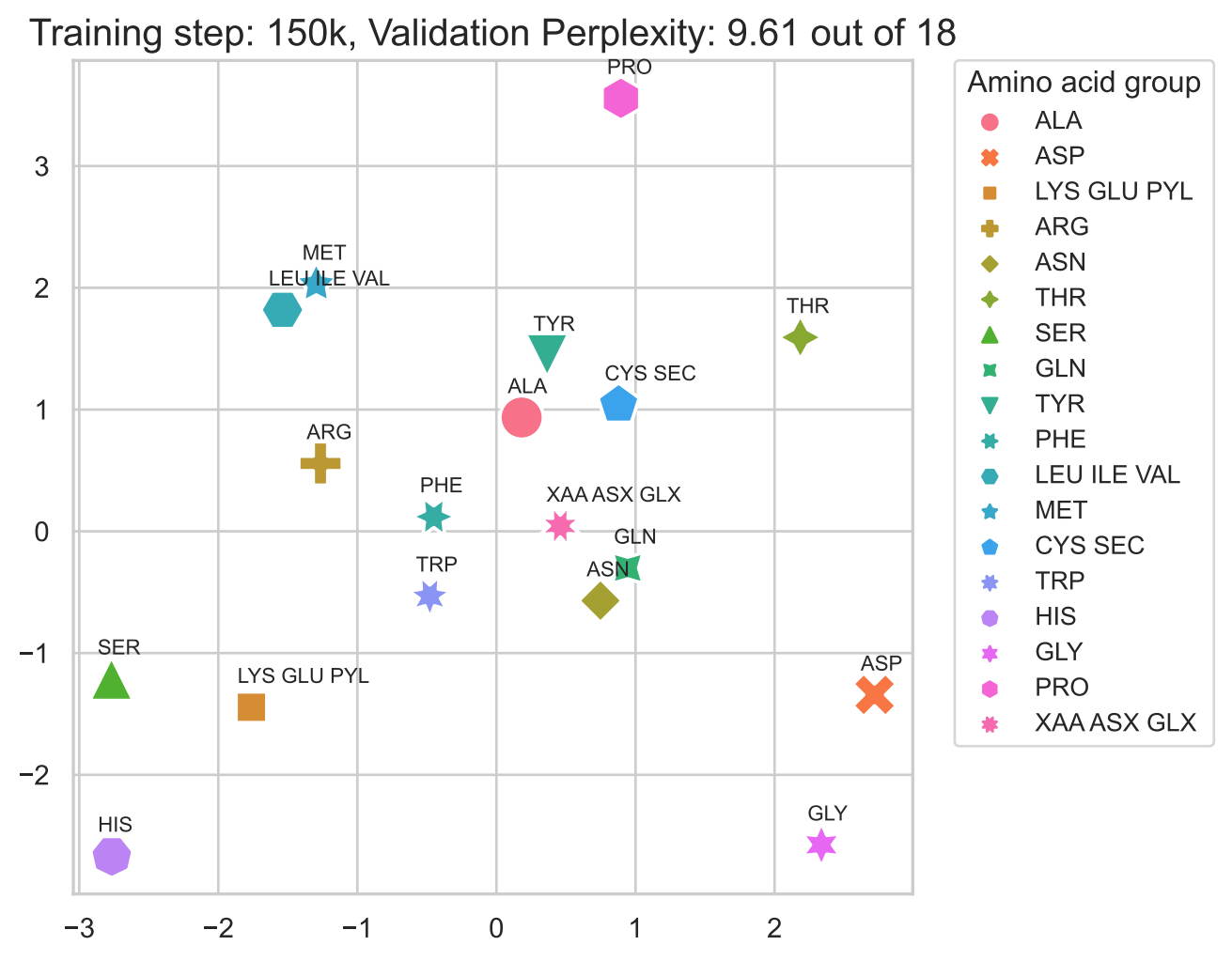

Training step: 100k, Validation Perplexity: 9.91 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
XAA ASX GLX
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

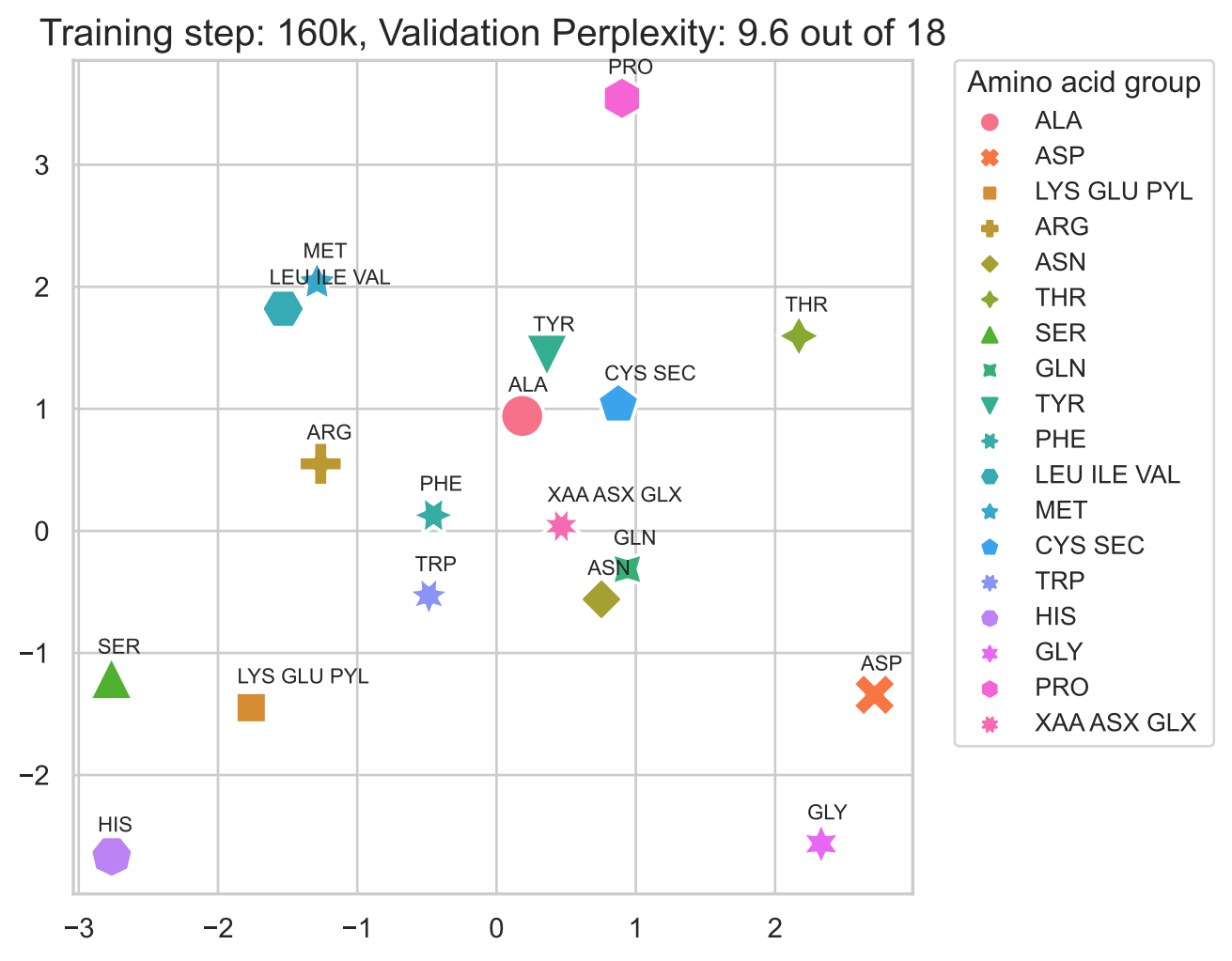

Training step: 110k, Validation Perplexity: 9.86 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
ARG
CYS SEC
PHE
XAA ASX GLX
ASN
GLN
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

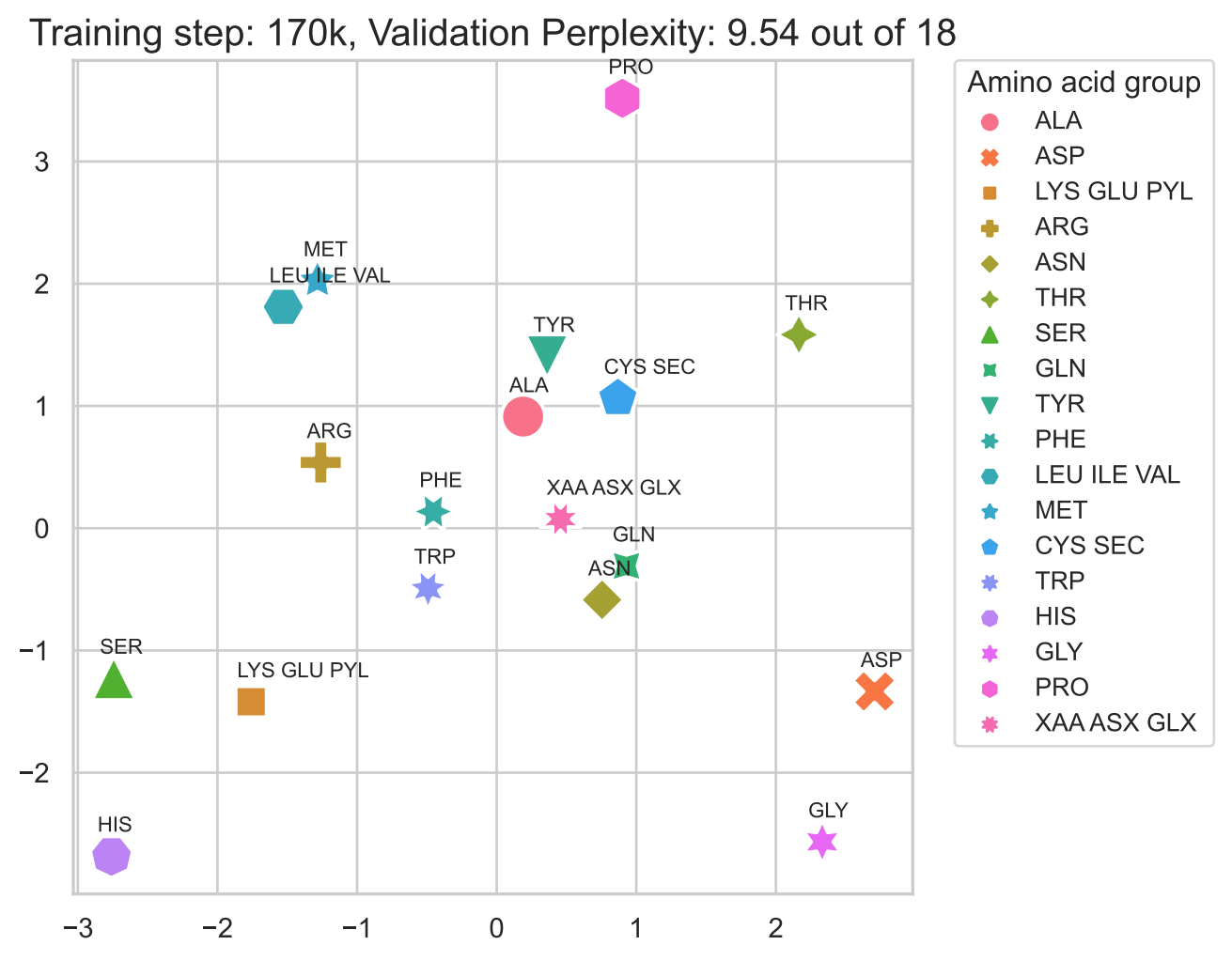

Training step: 120k, Validation Perplexity: 9.78 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
CYS SEC
ARG
PHE
XAA ASX GLX
ASN
GLN
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

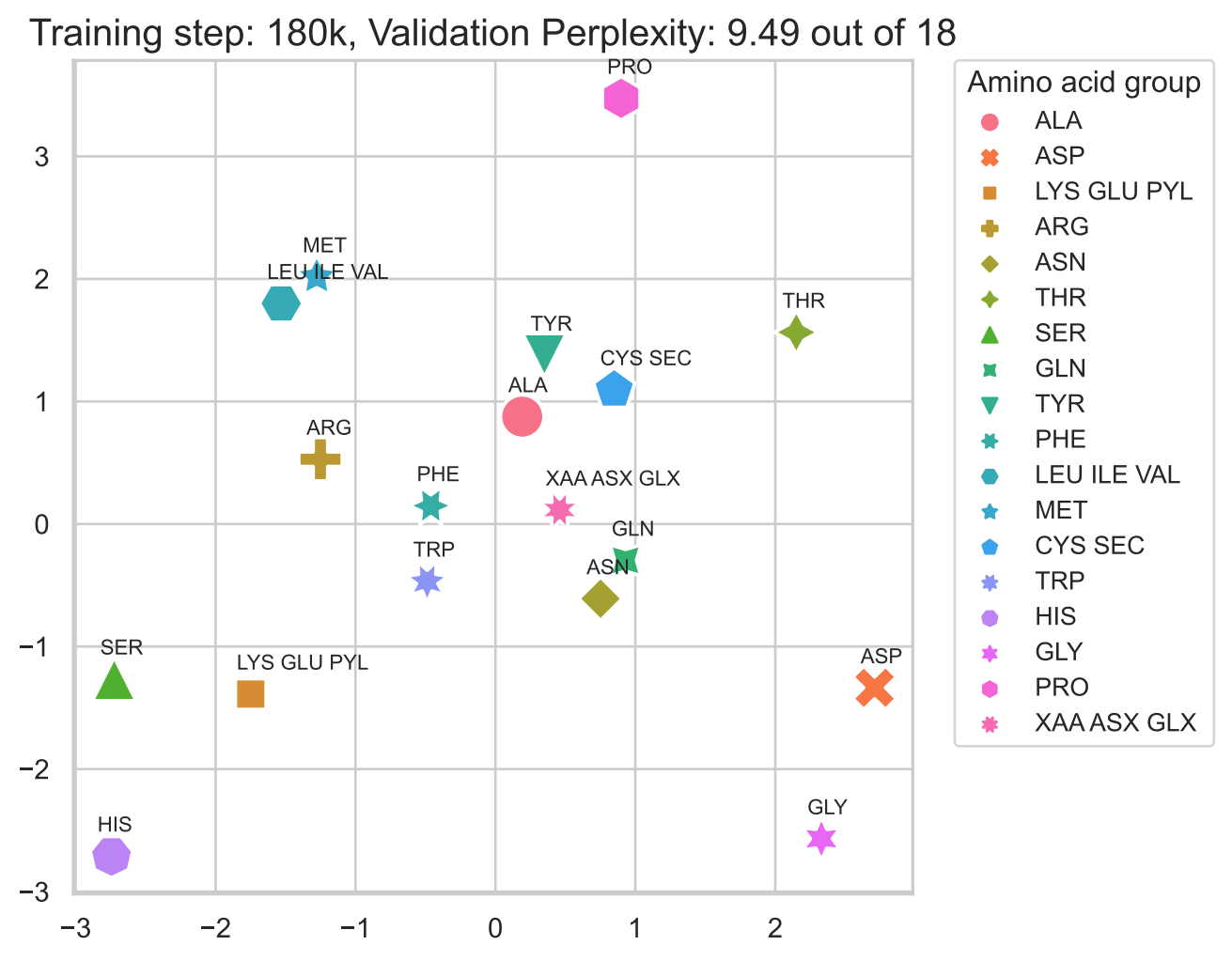

Training step: 130k, Validation Perplexity: 9.72 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
CYS SEC
ARG
PHE
XAA ASX GLX
GLN
ASN
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

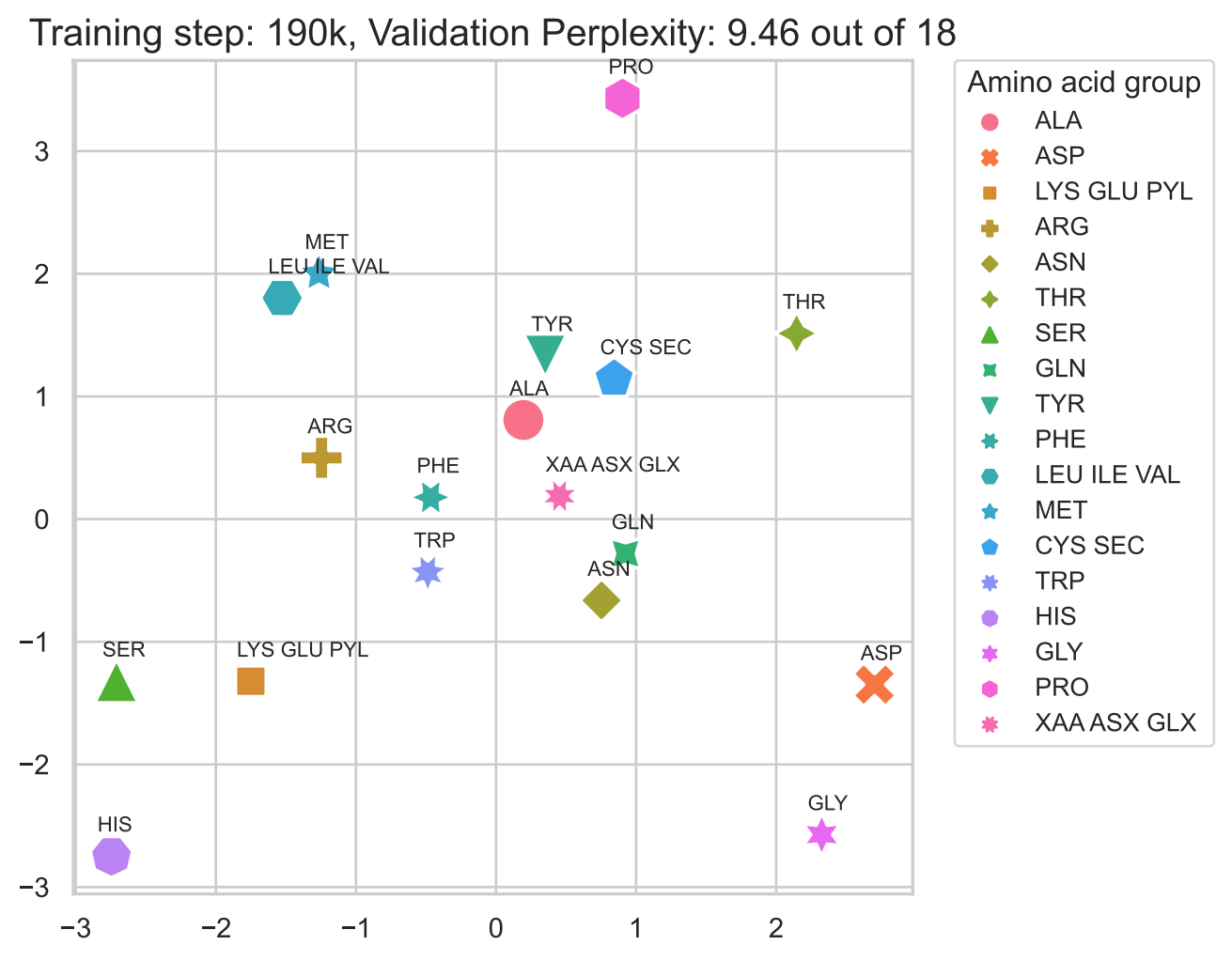

Training step: 140k, Validation Perplexity: 9.67 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
CYS SEC
ARG
PHE
XAA ASX GLX
GLN
ASN
TRP
SER
ASP
LYS GLU PYL
HIS
GLY
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

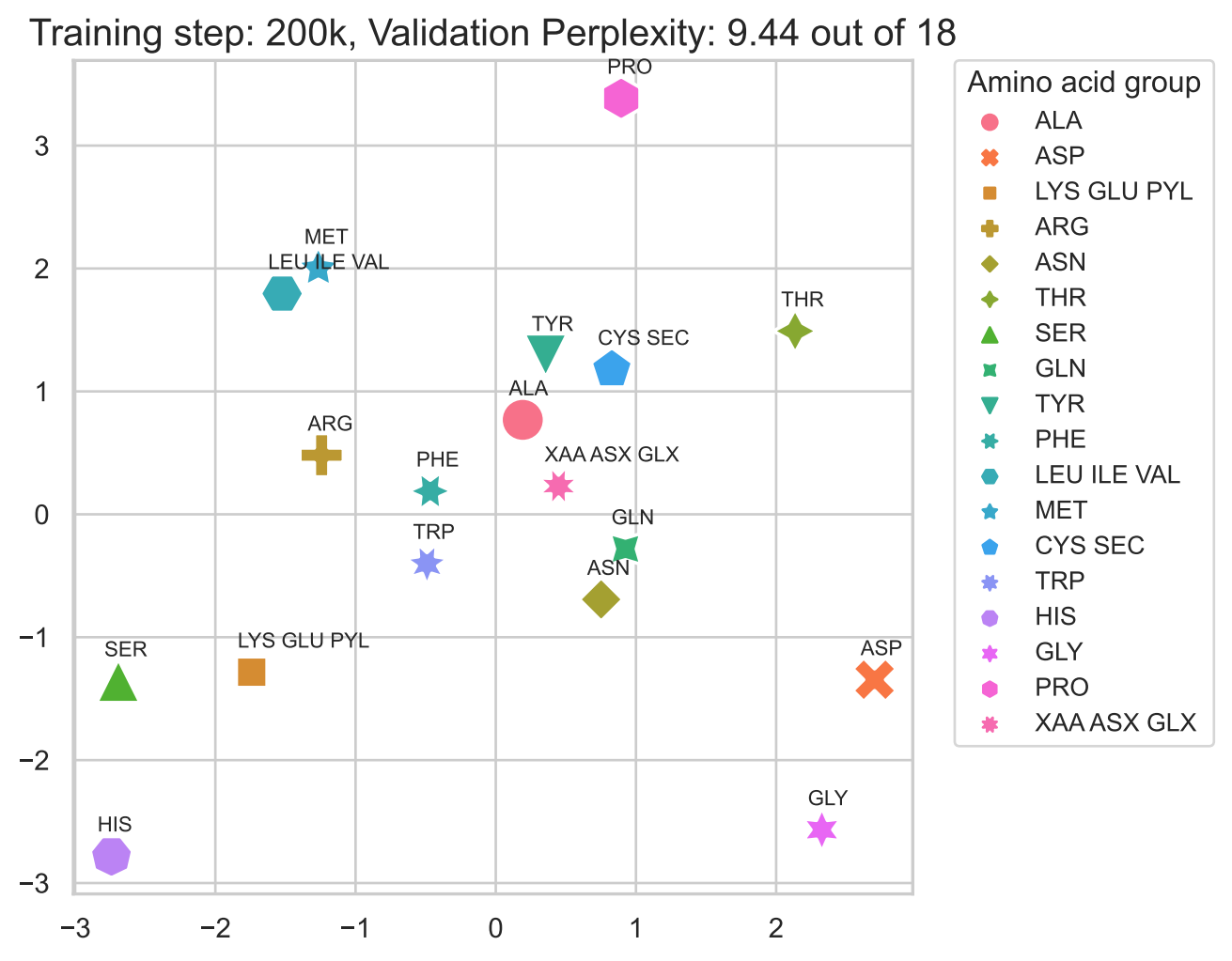

Training step: 150k, Validation Perplexity: 9.61 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
CYS SEC
ARG
PHE
XAA ASX GLX
GLN
TRP
ASN
SER
LYS GLU PYL
ASP
HIS
GLY
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

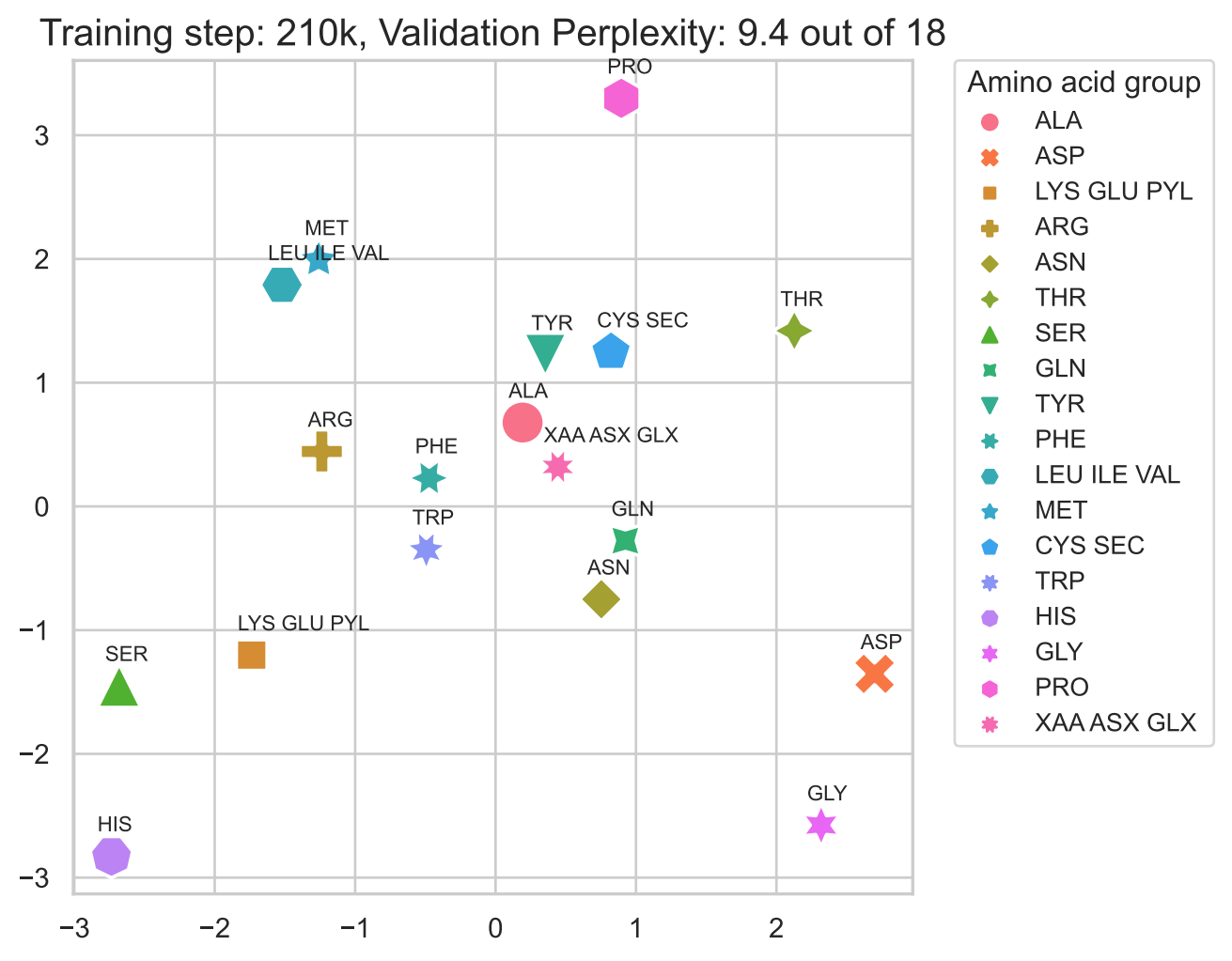

Training step: 160k, Validation Perplexity: 9.6 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
ALA
CYS SEC
ARG
PHE
XAA ASX GLX
GLN
ASN
TRP
SER
LYS GLU PYL
ASP
HIS
GLY
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−2
−1
0
1
2

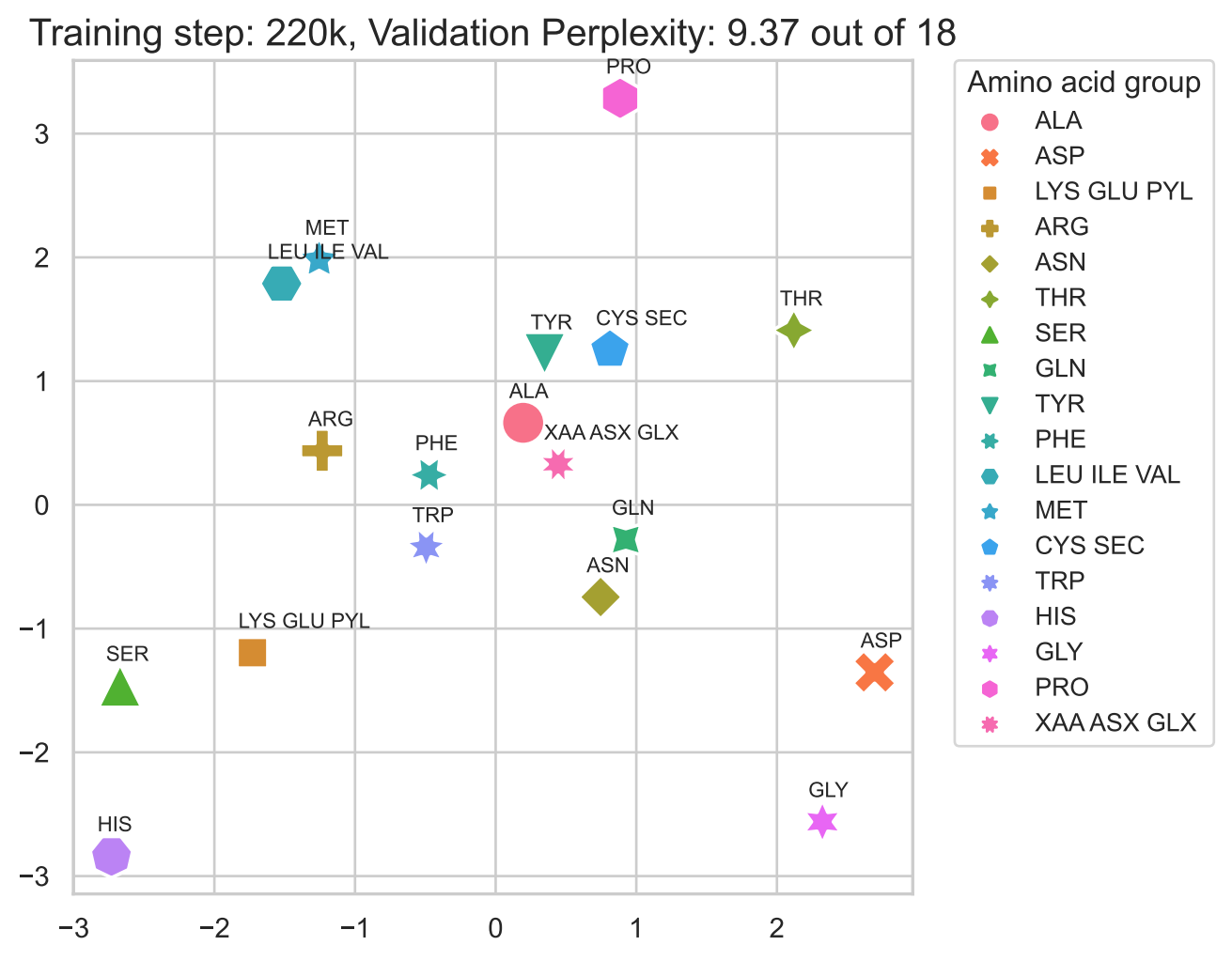

Training step: 170k, Validation Perplexity: 9.54 out of 18
PRO
MET
LEU ILE VAL
TYR
THR
CYS SEC
ALA
ARG
PHE
XAA ASX GLX
GLN
TRP
ASN
SER
LYS GLU PYL
ASP
HIS
GLY
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX

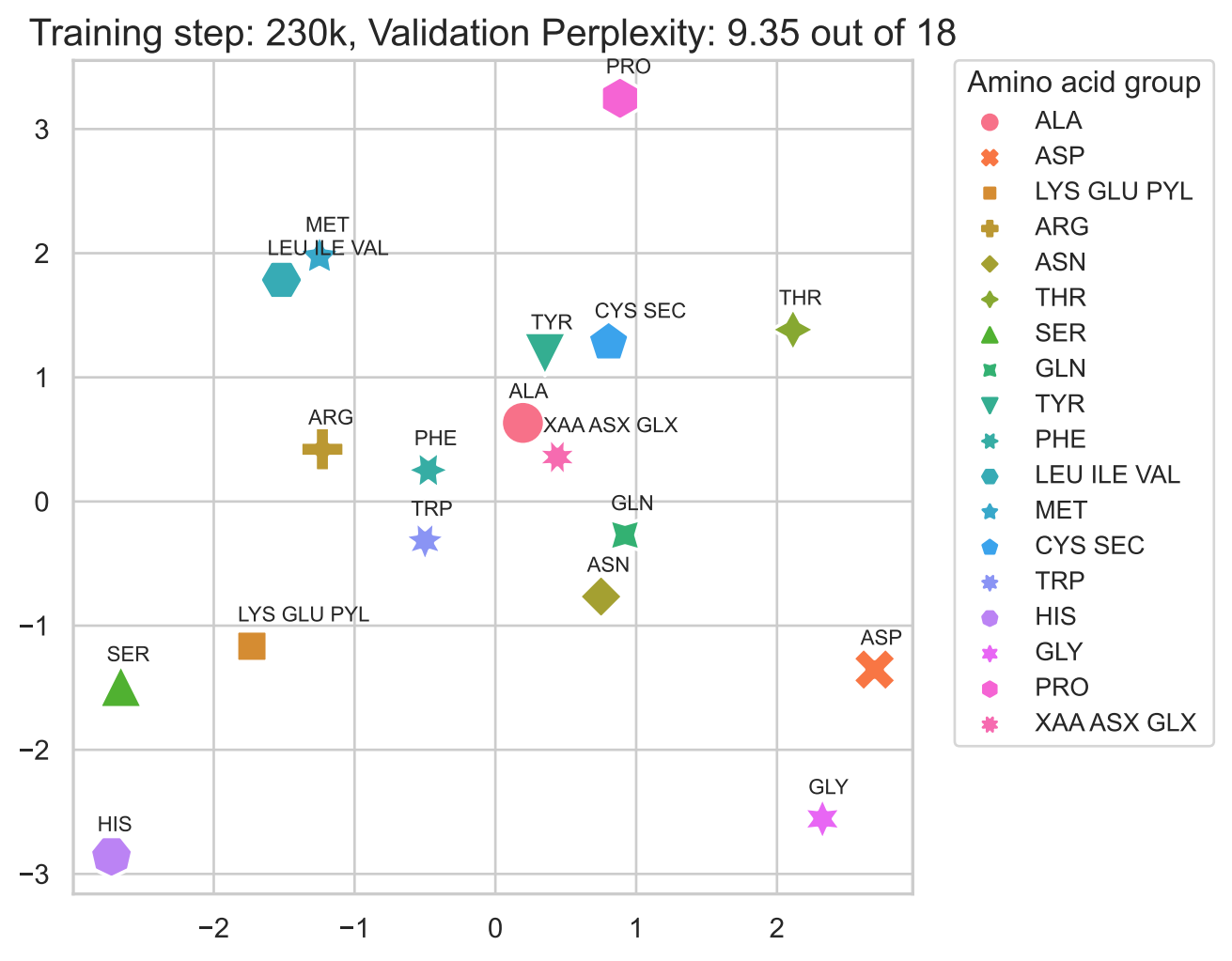

Training step: 180k, Validation Perplexity: 9.49 out of 18
PRO
MET
LEU ILE VAL
THR
TYR
CYS SEC
ALA
ARG
PHE
XAA ASX GLX
GLN
TRP
ASN
SER
LYS GLU PYL
ASP
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2

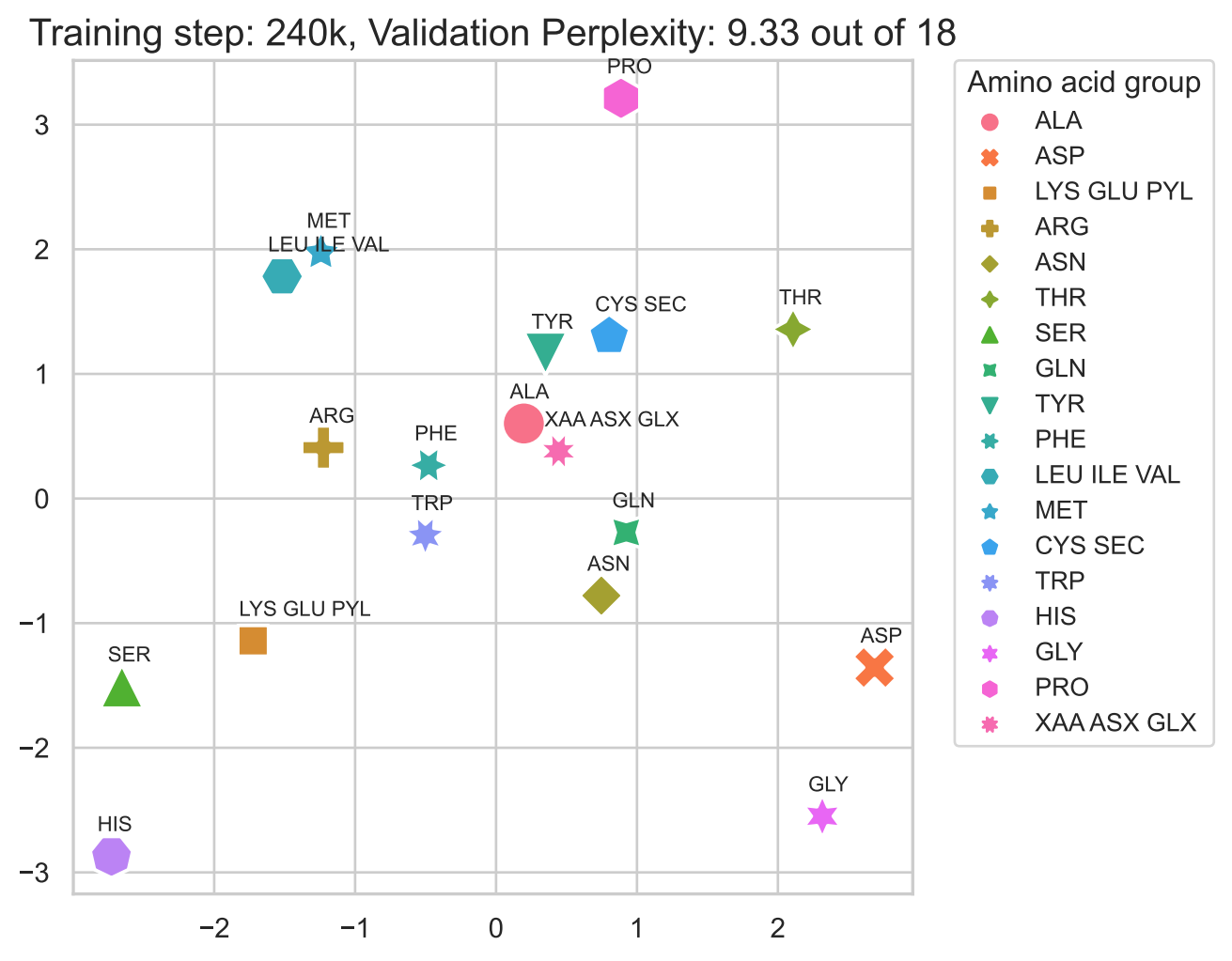

Training step: 190k, Validation Perplexity: 9.46 out of 18
PRO
MET
LEU ILE VAL
THR
TYR
CYS SEC
ALA
ARG
PHE
XAA ASX GLX
GLN
TRP
ASN
SER
LYS GLU PYL
ASP
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2


Training step: 200k, Validation Perplexity: 9.44 out of 18
PRO
MET
LEU ILE VAL
THR
TYR
CYS SEC
ALA
ARG
XAA ASX GLX
PHE
GLN
TRP
ASN
SER
LYS GLU PYL
ASP
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2


Training step: 210k, Validation Perplexity: 9.4 out of 18
PRO
MET
LEU ILE VAL
TYR
CYS SEC
THR
ALA
ARG
XAA ASX GLX
PHE
TRP
GLN
ASN
LYS GLU PYL
SER
ASP
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2


Training step: 220k, Validation Perplexity: 9.37 out of 18
PRO
MET
LEU ILE VAL
TYR
CYS SEC
THR
ALA
ARG
XAA ASX GLX
PHE
TRP
GLN
ASN
LYS GLU PYL
SER
ASP
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2


Training step: 230k, Validation Perplexity: 9.35 out of 18
PRO
MET
LEU ILE VAL
TYR
CYS SEC
THR
ALA
XAA ASX GLX
ARG
PHE
TRP
GLN
ASN
LYS GLU PYL
ASP
SER
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX
3
2
1
0
−1
−2
−3
−2
−1
0
1
2


Training step: 240k, Validation Perplexity: 9.33 out of 18
PRO
MET
LEU ILE VAL
CYS SEC
THR
TYR
ALA
XAA ASX GLX
ARG
PHE
TRP
GLN
ASN
LYS GLU PYL
ASP
SER
GLY
HIS
Amino acid group
ALA
ASP
LYS GLU PYL
ARG
ASN
THR
SER
GLN
TYR
PHE
LEU ILE VAL
MET
CYS SEC
TRP
HIS
GLY
PRO
XAA ASX GLX