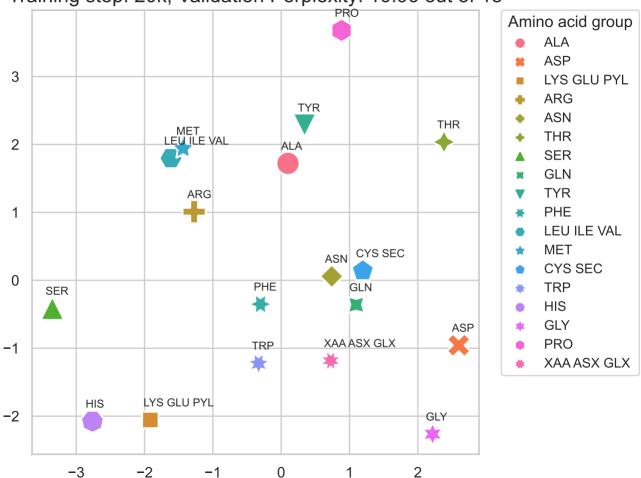
Training step: 10k, Validation Perplexity: 11.36 out of 18 Amino acid group ALA **ASP** 3 LYS GLU PYL **TYR ARG** THR ASN LEMETE VAL THR ALA 2 SER GLN **TYR** ARG PHE 1 LEU ILE VAL **MET** ASN CYS SEC CYS SEC 0 **TRP** GLN SER PHE HIS ASP **GLY PRO** -1 TRP XAA ASX GLX XAA ASX GLX GLY HIS LYS GLU PYL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

0

-3

Training step: 20k, Validation Perplexity: 10.96 out of 18



Training step: 30k, Validation Perplexity: 10.66 out of 18 Amino acid group ALA ASP 3 LYS GLU PYL **ARG TYR** ASN MET LEU ILE VAL THR THR 2 ALA SER GLN **TYR** ARG 1 PHE LEU ILE VAL CYS SEC MET ASN CYS SEC 0 PHE GLN **TRP** SER HIS **GLY** XAA ASX GLX ASP TRP -1 **PRO** XAA ASX GLX LYS GLU PYL HIS -2 GLY * 2 -3-20

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

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

-3

-2

2

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

Training step: 70k, Validation Perplexity: 10.56 out of 18 Amino acid group ALA **ASP** 3 LYS GLU PYL **ARG** ASN MET THR LEWILE VAL THR **TYR** 2 SER ALA GLN **TYR** ARG 1 PHE CYS SEC LEU ILE VAL MET PHE CYS SEC 0 **TRP** SER HIS TRP **GLY** * ASP **PRO** XAA ASX GLX LYS GLU PYL -2 HIS GLY 2 -3-2

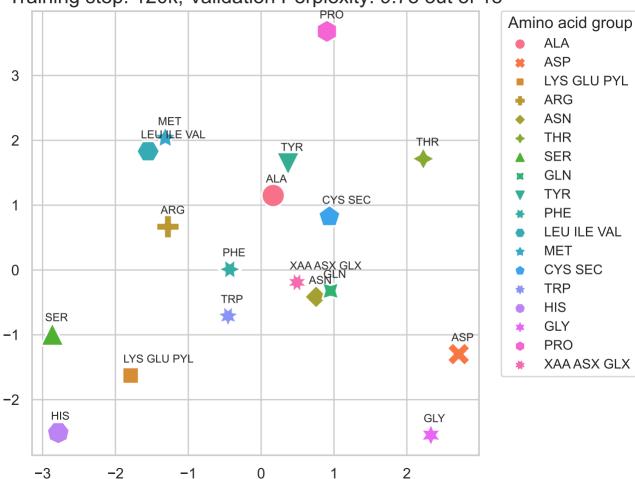
Training step: 80k, Validation Perplexity: 9.98 out of 18 Amino acid group ALA **ASP** 3 LYS GLU PYL **ARG** ASN MET THR LEUILE VAL THR **TYR** 2 SER ALA GLN **TYR** ARG 1 PHE CYS SEC LEU ILE VAL MET PHE CYS SEC 0 **TRP** SER HIS TRP **GLY** * ASP **PRO** XAA ASX GLX LYS GLU PYL -2 HIS **GLY** 2 -3-2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

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

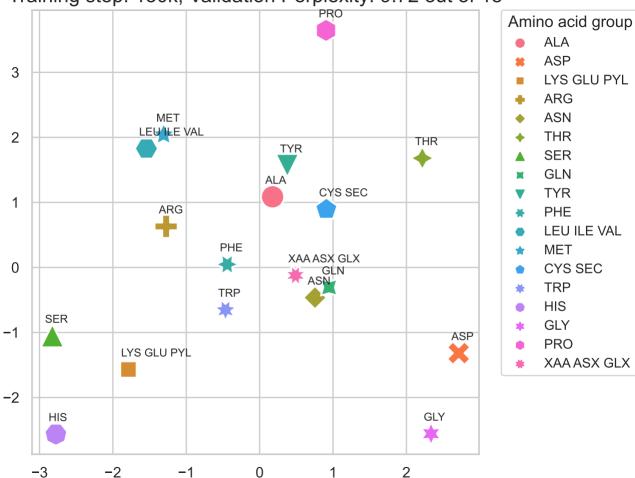
Training step: 100k, Validation Perplexity: 9.91 out of 18 Amino acid group ALA **ASP** 3 LYS GLU PYL **ARG** ASN MET THR LEUNLE VAL THR 2 **TYR SER** ALA GLN **TYR** 1 ARG CYS SEC PHE LEU ILE VAL **MET** PHE CYS SEC 0 XAAASISK IGLX **TRP** TRP HIS SER **GLY** -1 ASP **PRO** XAA ASX GLX LYS GLU PYL -2 HIS **GLY** 2 -3-2 0

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

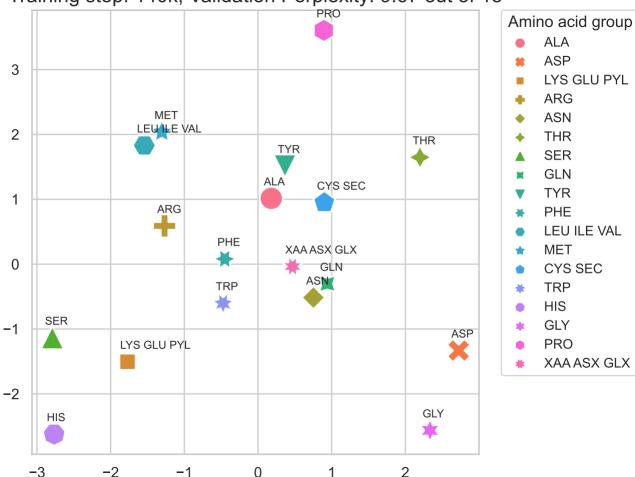
Training step: 120k, Validation Perplexity: 9.78 out of 18



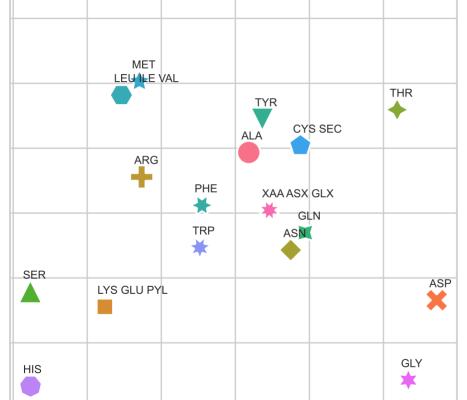
Training step: 130k, Validation Perplexity: 9.72 out of 18



Training step: 140k, Validation Perplexity: 9.67 out of 18



Training step: 150k, Validation Perplexity: 9.61 out of 18 PRO Amino acid group ALA **ASP** 3 LYS GLU PYL **ARG** MET ASN LEUNE VAL 2 THR THR **TYR SER** GLN CYS SEC ALA **TYR** 1 ARG PHE



0

0

-1

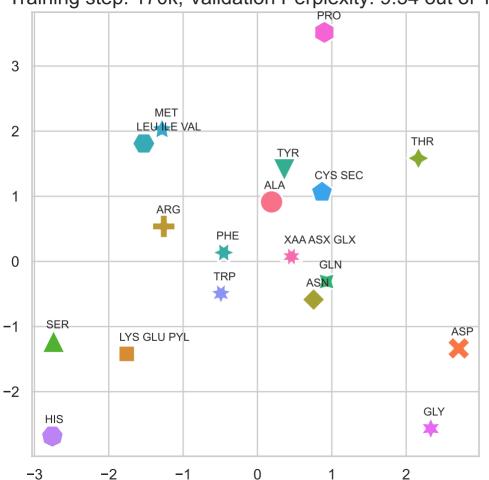
-2

-2



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

Training step: 170k, Validation Perplexity: 9.54 out of 18 PRO





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

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

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

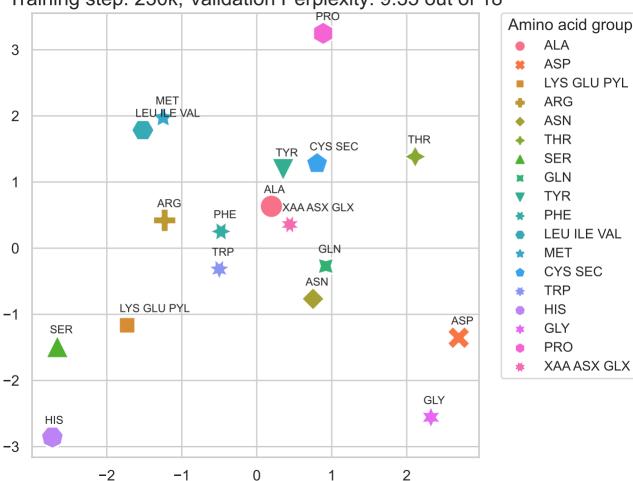
0

-2

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

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

Training step: 230k, Validation Perplexity: 9.35 out of 18 PRO Amino acid group ALA 3 **ASP** LYS GLU PYL MET **ARG**



Training step: 240k, Validation Perplexity: 9.33 out of 18 PRO Amino acid group 3 ALA ASP LYS GLU PYL MET **ARG** LEUILE VAL 2 ASN THR THR CYS SEC **TYR SER** GLN ALA **TYR** ARG XAA ASX GLX PHE PHE LEU ILE VAL 0 GLN TRP **MET** CYS SEC ASN **TRP** LYS GLU PYL HIS -1 ASP **GLY** SER **PRO** XAA ASX GLX -2 GLY HIS -3

0

-2

2