



**Figure 1: Scaling of the FIM perplexity with the number of context sequences.** We show the perplexity of the FIM part for different number  $N_m$  of masked amino acids versus the number of context sequences in a log-log plot. Results are averaged over all 500 clusters of the test set and 100 replicates for each cluster (differing by the random sampling of context sequences). Context sizes go up to  $2^{17}$  amino acids. To reduce noise, we take the exponential moving average, and we restrict to cases where the count of samples is at least 100.