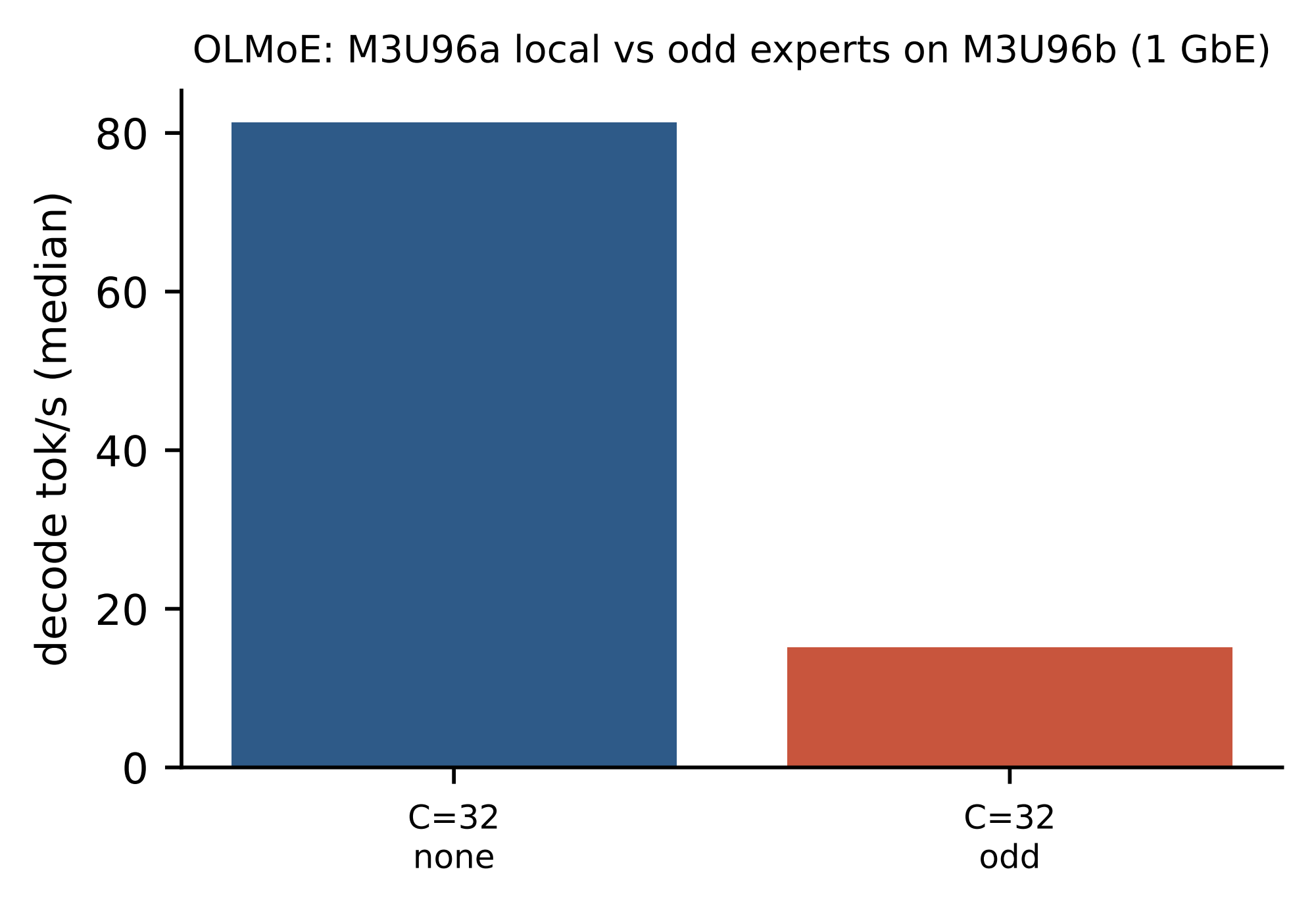

OLMoE: M3U96a local vs odd experts on M3U96b (1 GbE)
decode tok/s (median)
0
20
40
60
80
C=32
none
C=32
odd