ReinFlow-S (ours)   DPPO   FQL