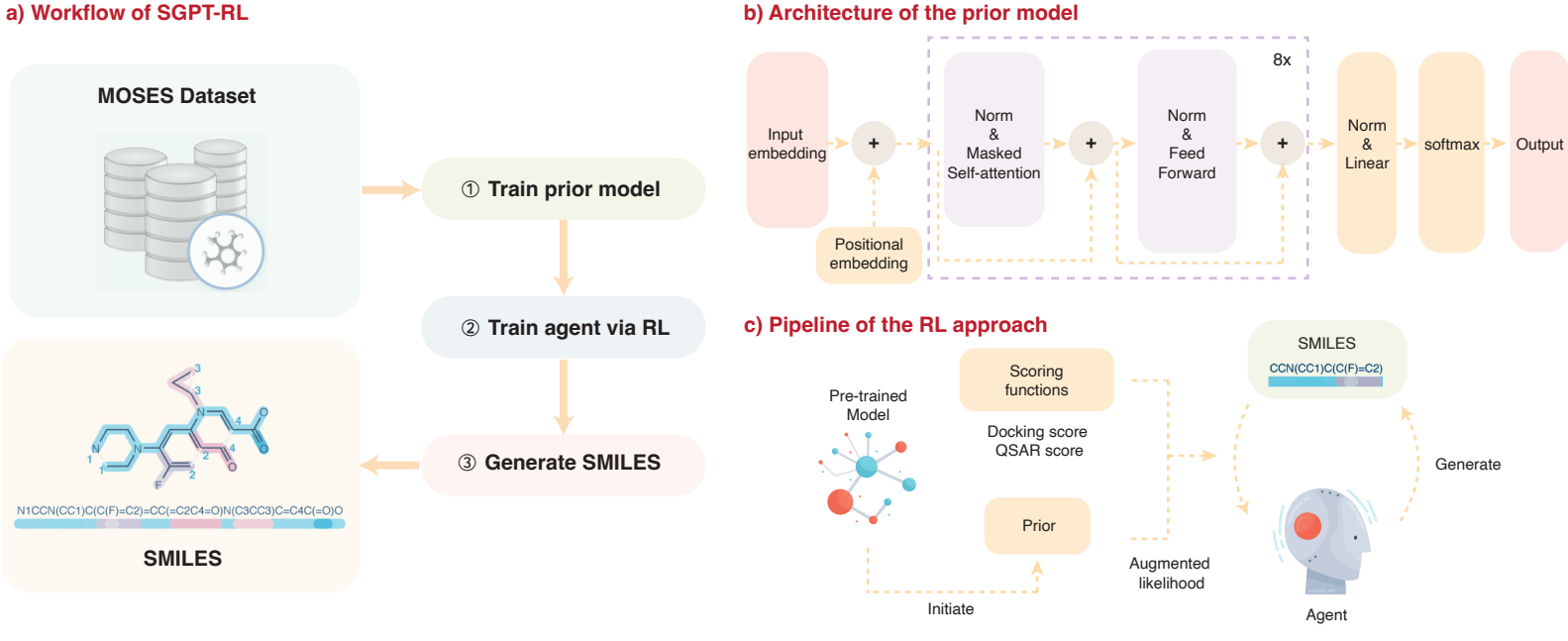
a) Workflow of SGPT-RL
MOSES Dataset
① Train prior model
② Train agent via RL
③ Generate SMILES
N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O
SMILES
b) Architecture of the prior model
Input embedding
Positional embedding
Norm & Masked Self-attention
Norm & Feed Forward
8x
Norm & Linear
softmax
Output
c) Pipeline of the RL approach
Pre-trained Model
Scoring functions
Docking score
QSAR score
Prior
Initiate
Augmented likelihood
SMILES
CCN(CC1)C(C(F)=C2)
Generate
Agent