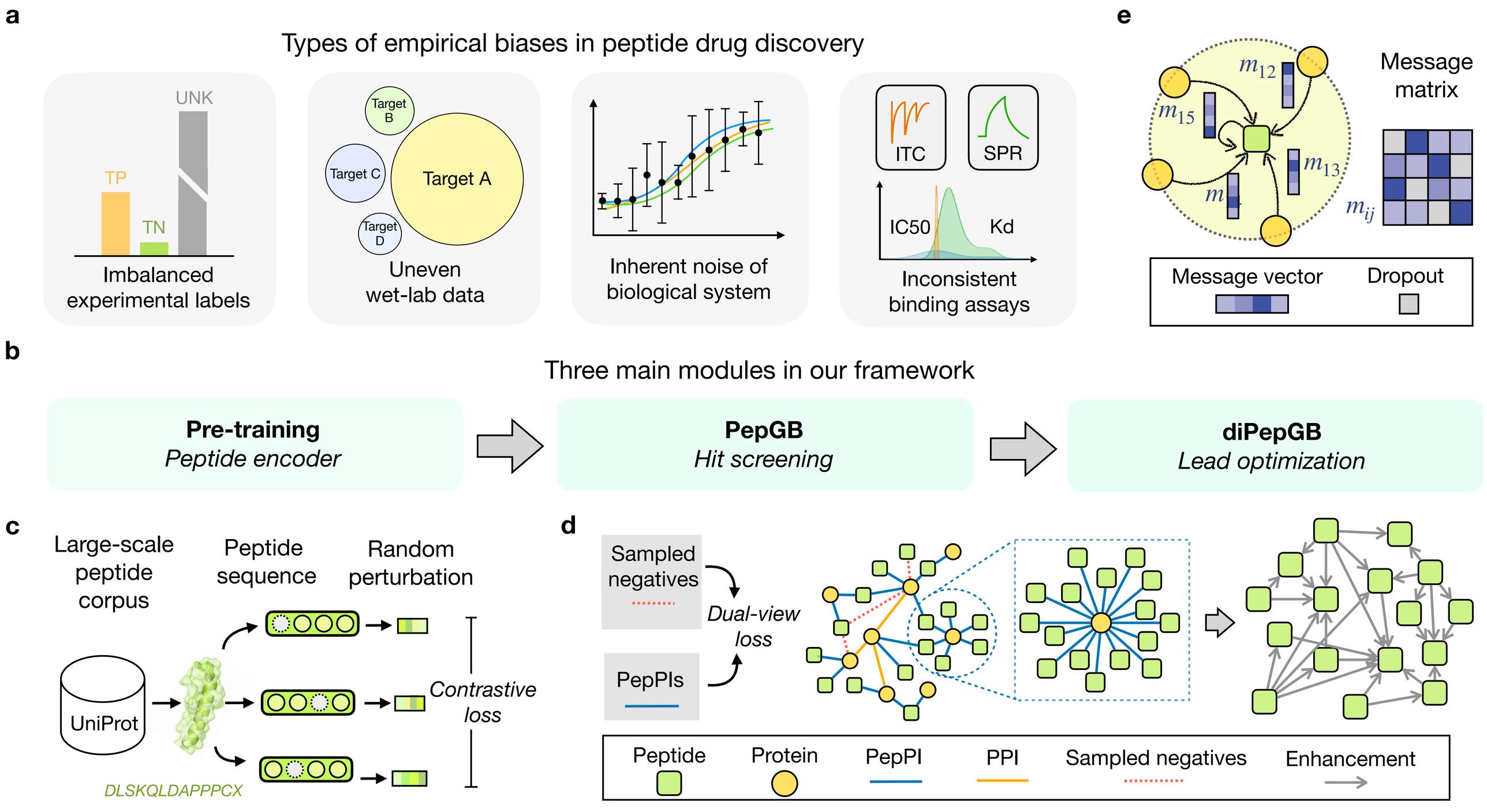
a
Types of empirical biases in peptide drug discovery
UNK
TP
TN
Imbalanced experimental labels
Target B
Target C
Target A
Target D
Uneven wet-lab data
Inherent noise of biological system
ITC
SPR
IC50
Kd
Inconsistent binding assays
e
m_{12}
m_{15}
m_{13}
m
m_{ij}
Message matrix
Message vector
Dropout
b
Three main modules in our framework
Pre-training
Peptide encoder
PepGB
Hit screening
diPepGB
Lead optimization
c
Large-scale peptide corpus
Peptide sequence
Random perturbation
UniProt
DLSKQLDAPPPCX
Contrastive loss
d
Sampled negatives
PepPIs
Dual-view loss
Peptide
Protein
PepPI
PPI
Sampled negatives
Enhancement