Proximity- and affinity- based labeling methods for interactome mapping

Beryl X. Li
March 13th, 2019
Proximity- and affinity-based labeling methods

Human protein-protein interactome contains a quarter million interactions between 22,000 proteins.

Protein-protein and protein ligand interactions elucidate biological processes and facilitate drug-target identification.

Classical methods
- yeast two-hybrid assay
- host incompatibility
- low throughput
- affinity-complex purification
- non-physiological conditions
- limited to high affinity interactions

Proximity labeling
- spatial resolution
- temporal resolution
- native environment

Proximity- and affinity-based labeling methods

general mechanism

bait
(protein or small molecule ligand)

prey
(target protein)

catalyst

tags near catalyst are activated

catalyst

isolate prey among a mixture of proteins

covalent linkage
Proximity- and affinity-based labeling methods

**Prevalent enzymatic methods**

- **BioID:** (Biotin IDentification)
  - **Split BioID**
  - **BioID2, BASU, TurboID**
- **APEX (Enhanced Ascorbate PeroXidase)**
  - **temporal resolution**


**Affinity-guided catalysts**

- **Ligand-tethered DMAP**
- **MoAL method (MOdular Affinity Labeling)**
  - **Local SET catalysis**
  - **Survey of photo-affinity labeling agents**

Proximity- and affinity-based labeling methods

Biotin IDentification (BioID)

Biotin ligase
R118G mutant to prematurely release biotinoyl AMP

Lamin-A (bait)
intermediate filament protein on nuclear lamina

prey

BirA*

biotinylation occurs in proximity to BirA* activity, which leads to selectivity

Proximity- and affinity-based labeling methods

Biotin IDentification (BioID)

Lamin-A (bait)
intermediate filament protein on nuclear lamina

BirA* is 35 kDa, might impact native binding
long incubation times lack temporal resolution
cells can express bait–ligase conjugate in vivo

identify SLAP75, novel nuclear envelope constituent

no predicted transmembrane domain
no clues in sequence motif

Proximity- and affinity-based labeling methods

BioID to map inhibitory postsynaptic density proteome

- inhibitory synapse dampen neuronal activity by postsynaptic hyperpolarization
- genetic perturbations strongly implicated in developmental brain disorders
  - molecular basis to synapse regulation is poorly understood

**Proximity- and affinity-based labeling methods**

**variations of BioID**

### Split-BioID

Ago protein plays at least two roles in gene-silencing

1) *represses mRNA in miRISC complex*

2) *load miRNA in RISC-loading complex*

how to assign novel identified proteins to a specific step?

---

### BioID2

- Biotin ligase from *Aquifex aeolicus* which naturally lacks N-terminus
  - 27 kDa (vs. original 35 kDa)

---

### BASU

- Biotin ligase from *Bacillus subtilis*
  - removing N-terminus did not impact activity; 28 kDa

---

### TurboID and MiniTurbo

- 14 and mutations from original BirA*
  - MiniTurbo is 28 kDa

- original efficient labeling in minutes instead of hours

- ideal combination of *temporal resolution* and *non-toxicity*

---

identified novel regulator *GIGYF2* with split BioID system, which previously AP/MS did not show association with Ago or TNRC6C

---


Proximity- and affinity-based labeling methods

APEX (Ascorbic peroxidase)

Ascorbic peroxidase (27 kDa) oxidizes biotin phenol in presence of \( \text{H}_2\text{O}_2 \)

- \( \sim 20 \text{ nm diffusion radius} \)
- \( \text{H}_2\text{O}_2 \) may be harsh for labeling conditions
- Great temporal resolution, labeling in 1 min

Proximity- and affinity-based labeling methods

APEX resolves GPCR networks in vivo

G-protein-coupled receptors (GPCRs) mediate physiological responses to many *stimuli* (e.g., hormones, neurotransmitters, light, etc.)

How to track the series of cascading protein-protein interactions following agonist binding?

![Diagram showing GPCR interactions over time](image)

Proximity- and affinity-based labeling methods

APEX resolves GPCR networks in vivo

G-protein-coupled receptors (GPCRs) mediates physiological responses to many *stimuli* (e.g., hormones, neurotransmitters, light, etc.)

How to track the series of cascading protein-protein interactions following agonist binding?

Proximity- and affinity-based labeling methods

acylation of lectins by ligand-tethered DMAP catalysts

Congerin II (animal lectin with high lactose affinity)

- sugars-binding proteins (e.g., lectins) play important roles in glycobiology
- sugar-lectin interactions not sufficiently strong for affinity purification

use DMAP as acyl-transfer catalyst to covalently tag sugar-binding proteins

Proximity- and affinity-based labeling methods

acylation of lectins by ligand-tethered DMAP catalysts

35% yield in 3 h

Proximity- and affinity-based labeling methods

acylation of lectins by ligand-tethered DMAP catalysts


35% yield in 3 h
single product

lactose (ligand) fluorescein (acyl donor)
Proximity- and affinity-based labeling methods

“multivalent” DMAP catalysts for live-cell imaging

bradykinin $B_2$ (ligand)

[Chemical structure image]

- selectively label bradykinin $B_2$ receptor on cell surface
- development of fluorescent biosensor to screen potential antagonist ligand binding

Proximity- and affinity-based labeling methods

modular affinity labeling based on catalytic amidation

prey (target protein)

activatable tag with reactive module

bait (protein or small molecule ligand)

catalyst

catalyst

covalent linkage

possibly synthetically challenging

ligand activity may alter due to modification
Proximity- and affinity-based labeling methods

modular affinity labeling based on catalytic amidation

prey (target protein)

reactive module

activatable tag

bait (protein or small molecule ligand)

catalyst

possibly synthetically challenging

ligand activity may alter due to modification

modular approach, where ligand, tags, and reactive module are completely separated

Proximity- and affinity-based labeling methods

modular affinity labeling based on catalytic amidation

169% yield after 8 h (four binding sites/avidin)

negative control gave 11% yield

Proximity- and affinity-based labeling methods

photoredox proximity-based labeling


Proximity- and affinity-based labeling methods

photoredox proximity-based labeling


selectively label carbonic anhydrase in mouse erythrocyte lysate
**Proximity- and affinity-based labeling methods**

*photo-affinity labeling*

**Advantages**
- reactive under specific activation conditions
- (potentially) mild conditions
- covalently modify targets for easy purification

**Applications**
- identification of membrane protein targets
- elucidation of protein structure in solution
- characterization of proteins in pharmaceutical solids

**Nitrenes**
- insertion into C–H, N–H, or O–H bonds
- ring-expansion with nucleophilic addition
- preference for cysteine and aromatic residues

\[
\text{N}_3 \xrightarrow{h\nu} \text{N}^+\cdot
\]

\[
\text{N}_3
\]

\[
260 \text{ nm}
\]

**Carbenes**
- insertion into C–C and X–H bonds (X = C, O, N, S)
- addition to C=C bonds
- preference for cysteine and aromatic residues

\[
\text{CF}_3\text{N}^+\cdot
\]

\[
\text{CF}_3\text{N}^-
\]

\[
330-370 \text{ nm}
\]

**Benzophenone**
- triplet ketyl biradical
- high affinity towards methionine
- relatively long wavelength activation: less damaging to proteins

\[
\text{O}^+\cdot
\]

\[
\text{O}^-
\]

\[
365 \text{ nm}
\]