**7AUY-Eosin-Y protocol**

Since the active site engineering worked well with 7AUY and mes-acr, start with this method first.

1a. Diffuse residues within 8 angstroms of substrate:

['9', '10', '12', '13', '14', '16', '18', '22', '23', '24', '26', '27', '28', '29', '30', '31', '32', '33', '34', '35', '72', '73', '75', '76', '77', '78', '79', '80', '81', '82', '85', '86', '89', '106', '108', '142', '143', '144', '145', '146', '171', '172', '198', '199', '201', '202', '238', '266', '267', '268', '269', '270', '271', '272', '273']

Command:

*python run\_inference.py inference.deterministic=True diffuser.T=200 inference.output\_prefix=output/ligand\_protein\_motif/7auy\_eos\_8A\_ inference.input\_pdb=input/7auy\_EOS.pdb contigmap.contigs=[\'1-10,A2-8,20-35,A36-45,1-5,A48-58,1-3,A60-67,1-3,A69-71,10-20,A90-105,10-20,A120-141,8-12,A150-170,25-40,A203-216,1-3,A218-265,10-50\'] contigmap.length="250-350" inference.ligand=EOS inference.num\_designs=1 inference.design\_startnum=0*

1b. Diffuse residues within 5 angstroms of substrate:

[13, 27, 30, 31, 34, 75, 76, 80, 81, 86, 143, 144, 202, 267, 268, 269, 270, 273]

Command:

*python run\_inference.py inference.deterministic=True diffuser.T=200 inference.output\_prefix=output/ligand\_protein\_motif/7auy\_eos\_5A\_ inference.input\_pdb=input/7auy\_EOS.pdb contigmap.contigs=[\'A2-12,2-5,A14-26,5-12,A35-45,1-3,A48-58,1-3,A60-67,1-3,A69-73,10-18,A87-108,8-15,A120-142,2-4,A145-179,15-25,A198-201,1-3,A203-216,1-3,A218-265,10-50\'] contigmap.length="250-350" inference.ligand=EOS inference.num\_designs=1 inference.design\_startnum=0*

2. Protein MPNN then alphafold3 for prediction

a) Select 2 best aligned from 5 and 8 for protein MPNN, 4 structures for protein MPNN then align with diffused structure.

- 5A: diffused structures 0 and 4 chosen

- 8A: diffused structure 3 and 4 chosen

\*\*\*\*\*\*\*

5A results:

structure 0 - protein mpnn sequence id4 and id14 chosen.

Results of alignment with template (3/6/2024):

id4: RMSD 3.324A, obvious steric clash with substrate, REJECTED.

id14: RMSD 5.394A, possible steric clash, **try next step** and see if ligand mpnn improves it.

structure 4 - protein mpnn sequence id7 and id15 chosen.

Results of alignment with template (3/6/2024):

id7: RMSD 5.944, possible steric clash, REJECTED.

id15: RMSD 4.957, pocket looks promising, **move to next step**.

----------------------------------------------

8A results:

structure 3 - protein mpnn sequence id5 and id14 chosen.

Results of alignment with template (3/6/2024):

id5: RMSD 3.021, very likely steric clash, pocket looks too small, REJECTED.

id14: RMSD 4.559, some steric clash, very large pocket, **move to next step**.

structure 4 - protein mpnn sequence id13 and id14 chosen.

Results of alignment with template (3/6/2024):

id13: RMSD 1.746, promising pocket, **move to next step**.

id14: RMSD 2.528, promising pocket, **move to next step**.

\*\*\*\*\*\*\*

b) Also do protein MPNN only for 7auy\_EOS template, 1 structure for protein MPNN

\*\*\*\*\*\*\*

7AUY-EOS structure – protein mpnn sequence id5 and id12 chosen.

Results of alignment with template (3/6/2024):

id5: RMSD 1.976, promising pocket, **move to next step**.

id12: RMSD 1.696, promsing pocket, slight chance of clash, **move to next step**.

\*\*\*\*\*\*\*

3. Ligand MPNN of the best aligned structures using 3 different ligand MPNN parameters for each structure:

1. Normal
2. Higher temperature: 0.25
3. Side-chain packing

\*\*\*\*\*\*\*

Ligand MPNN results 4/6/2024

1 sequence chosen for each parameter by overall confidence score, see excel for chosen sequences and score. 3 sequences per structure, total of 21 sequences including protein mpnn only design.

\*\*\*\*\*\*\*