

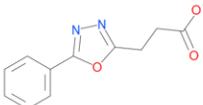
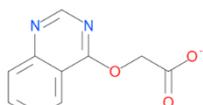
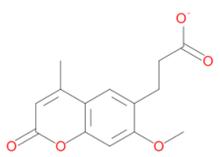
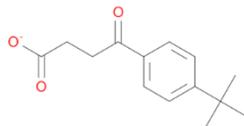
Co-crystal structures of USP5 Zf-UBD and weak binding compounds

Objective: To find crystallization conditions which allow growth of well-diffracting co-crystals of USP5 zinc finger ubiquitin binding domain (Zf-UBD) and compounds shown to bind weakly by **19F NMR** and **SPR assays** & to solve the co-crystal structures to determine if electron density of ligands can be seen in the binding pocket of the protein domain.

Experiment & Results:

Based on previous 19F NMR and SPR assay experiments, the compounds listed in Table 1 were used to set up co-crystallization screen with USP5¹⁷¹⁻²⁹⁰. 300 µL of 1:5 and 1:2.5 protein:compound solutions were prepared for each of the compounds in Table 1 in 50 mM Tris pH 8, 150 mM NaCl, 1 mM TCEP, where the concentration of USP5¹⁷¹⁻²⁹⁰ was constant at 12 mg/mL.

Table 1. USP5 Zf-UBD ligands

Compound Name	Compound Structure	SMILES
DAT180		<chem>C(Cc1nnc(c2ccccc2)o1)C([O-])=O</chem>
DAT194		<chem>C(C([O-])=O)Oc1c2ccccc2ncn1</chem>
DAT198		<chem>Cc1cc(=O)oc2c1cc(c(c2)OC)CCC(=O)O</chem>
DAT201		<chem>CC(C)(C)c1ccc(cc1)C(CCC([O-])=O)=O</chem>

Using SGC standard crystallization protocols, **SGC and RW screen** were used to prepare a crystal screen in a 96-well Intelli plates (Art Robbins Instruments). 70 µL of each condition was dispensed into the well of the plate and then 0.5 µL of the well solution was dispensed to both drops in the crystal plate by a liquid handling robot (Phoenix) followed by 0.5 µL of 1:5 and 1:2.5 protein:compound solutions in drop 1 (top drop) and drop 2 (bottom drop) respectively. Crystal plates were sealed and stored at 18°C.

Five days after preparing the crystal plates, crystals formed for some of the USP5-compound complexes. These results are summarized in Table 2. Please refer to the SGC and RW screens attached for components of crystallization conditions.

Table 2. Co-crystals of USP5¹⁷¹⁻²⁹⁰ and compounds

Compound Name	Crystallization Condition
DAT180	- SGC screen: B12, C04, E04, F04, H03, H12 - RW screen: B01, C01, F07
DAT194	- SGC screen: H03, H12, B12
DAT198	- No crystals formed
DAT201	- SGC screen: F03, G03, H03, H04, F04, E04, C04, B04, A04, H12 - RW screen: B01, C01

The largest and most 3D crystals for each compound were mounted using a nylon loop then transferred to a 2 μ L drop of well solution supplemented with 25% ethylene glycol (v/v) and submerged for approximately 30 seconds, and then cryo-cooled in liquid nitrogen. The crystal was screened using our in house diffractometer, RIGAKU FR-E SUPERBRIGHT at 1.54178 Å. 2 images at 90 degrees with a 0.5 degrees oscillation, 20 s exposure and 100 mm crystal-detector distance were collected with a RIGAKU SATURN A200 CCD detector at 100 K. All crystal conditions for the compounds had an oily skin on the top of the drop which needed to be removed prior to crystal mounting and began to melt slightly when cryo-protected with ethylene glycol which is why the quick cryo-soak protocol was used. Figures 1-3 are diffraction images with the crystals. No ice rings are visible and diffraction is better than 2 Å resolution for screened crystals.

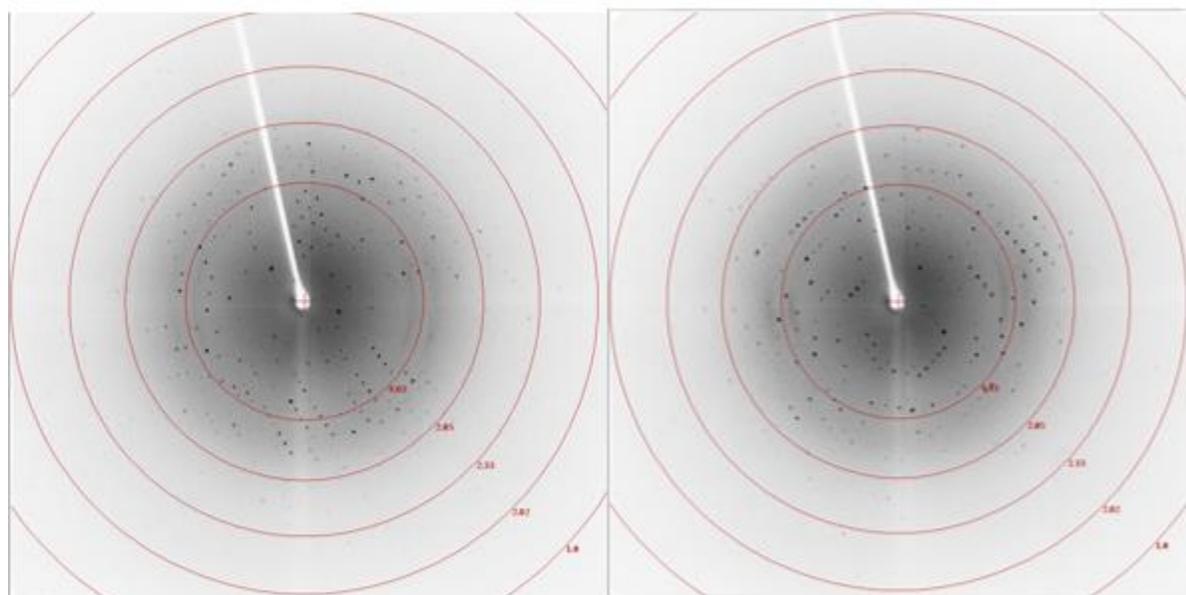


Figure 1. 1:2.5 USP5¹⁷¹⁻²⁹⁰: DAT180 in RW B01 (1.5 M ammonium sulfate, 0.1 M bis-tris pH 7.0, 1.1% DMSO (v/v), 25% ethylene glycol (v/v))

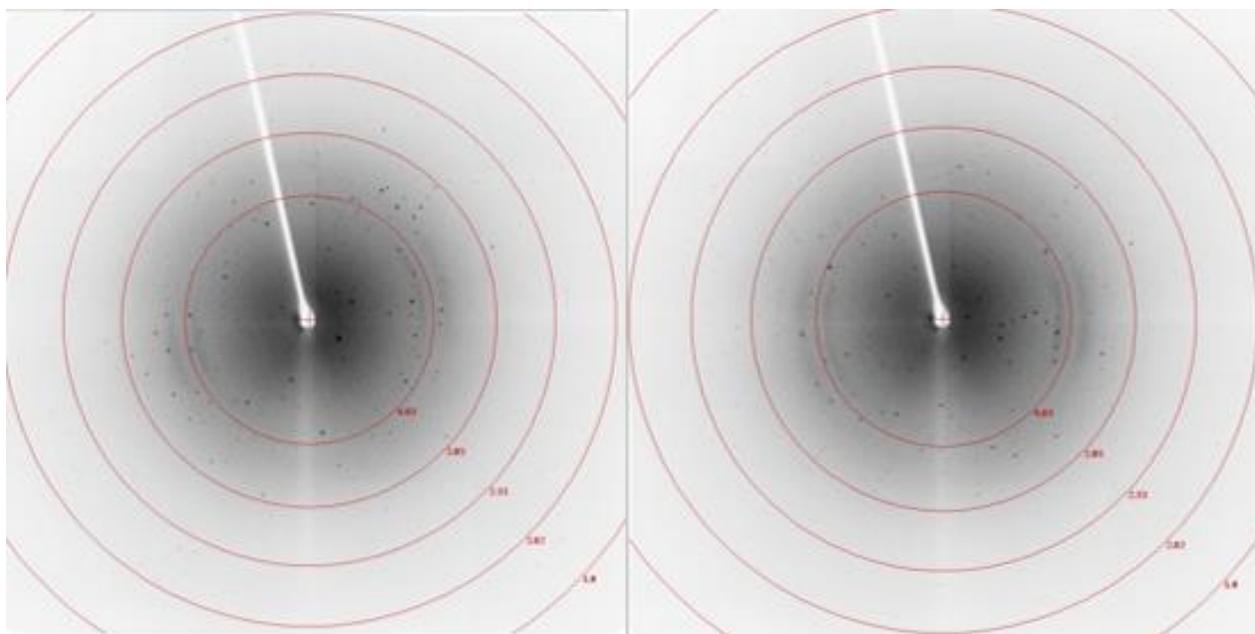


Figure 2. 1:2.5 USP5¹⁷¹⁻²⁹⁰:DAT194 in SGC H12 (2 M sodium/potassium phosphate pH 7.0, 1.1% DMSO (v/v), 25% ethylene glycol (v/v))

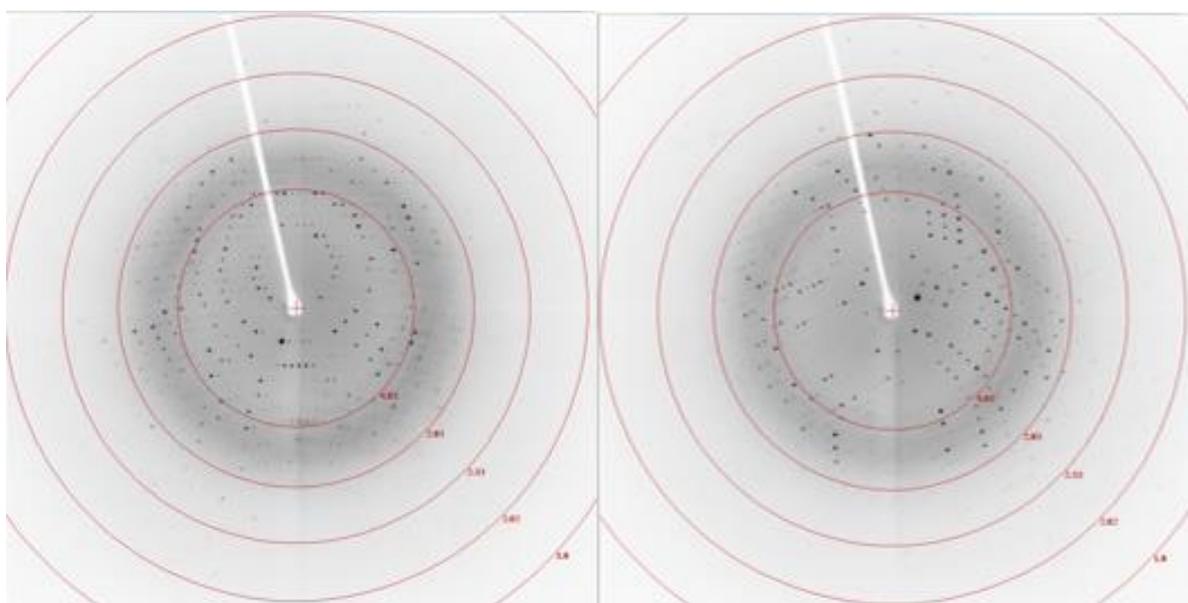


Figure 3. 1:5 USP5¹⁷¹⁻²⁹⁰: DAT201 RW B01 (1.5 M ammonium sulfate, 0.1 M bis-tris pH 7.0, 2.25% DMSO (v/v), 25% ethylene glycol (v/v))

I was able to grow well-diffracting crystals of USP5¹⁷¹⁻²⁹⁰ co-crystallized with DAT180, a DAT194 isomer and DAT201. DAT180 was co-crystallized with USP5¹⁷¹⁻²⁹⁰ in 1.5 M ammonium sulfate, 0.1 M bis-tris pH 7.0, 1.1% DMSO (v/v), 25% ethylene glycol (v/v). DAT194 was co-crystallized with USP5¹⁷¹⁻²⁹⁰ in 2 M sodium/potassium phosphate pH 7.0, 1.1% DMSO (v/v), 25% ethylene glycol (v/v). DAT201 was co-crystallized with USP5¹⁷¹⁻²⁹⁰ in 1.5 M ammonium sulfate, 0.1 M bis-tris pH 7.0, 2.25% DMSO (v/v), 25% ethylene glycol (v/v).

Large scale data collection of the crystals from Figures 1-3 were collected using the following conditions: crystal-detector distance: 60 mm, 180 images, 1 degree oscillation and an exposure time of 10 seconds.

DAT180 and 194 Images were processed with Xia2, scaled with AIMLESS. The space group was shown to be the same as existing 2G43 so the structure could be solved with direct refinement with the complete 2G43 PDB structure (chains A and B). DAT201 images were processed with Xia2, scaled with AIMLESS and phased using Phaser with chain A of PDB: 2G43 as the starting model.

All structures were then refined with iterative builds in Coot and refinement in Refmac. Clear electron density was seen in the binding pocket for each of the datasets. The resulting pdb, aimless mtz files, ligand cif files, omit maps files, symmetry coordinates and mmcif files are attached for each compound.

It is to be noted that DAT180 and DAT201 structures are refined and have already been deposited in the PDB (6DXT, 6DXH) and are awaiting release. The DAT194 structure is not processed or refined to the point of being ready for PDB deposition and should be interpreted with caution. Table 3 summarizes the solved co-crystal structures for each compound as well as previously deposited apo (2G43) and ubiquitin bound (2G45) structures.

Table 3. Summary of co-crystal structures

Compound Name	Rfactor/Rfree	Symmetry: Space group	Unit cell		Resolution (Å)
			Length (Å)	Angle (°)	
PDB: 2G43	0.23/0.27	C 1 2 1	a=61.38 b=86.18 c=59.9	$\alpha=90.00$ $\beta=99.29$ $\gamma=90.00$	2.09
PDB: 2G45	0.23/0.27	P 64	a=68.07 b=68.07 c=225.35	$\alpha=90.00$ $\beta=90.00$ $\gamma=120.00$	1.99
DAT180 (PDB: 6DXT)	0.19/0.24	C 1 2 1	a=61.09 b= 85.44 c=59.74	$\alpha=90.00$ $\beta=100.29$ $\gamma=90.00$	1.95
DAT194	0.18/0.21	C 1 2 1	a=61.87 b= 85.04 c=59.83	$\alpha=90.00$ $\beta=99.00$ $\gamma=90.00$	1.62
DAT201 (PDB: 6DXH)	0.21/0.25	I 2 2 2	a=47.69 b= 81.50 c=99.66	$\alpha=90.00$ $\beta=90.00$ $\gamma=90.00$	2.0

DAT180

Electron density of DAT180 was clearly seen in the binding pocket of USP5¹⁷¹⁻²⁹⁰ (chain A) and could be modeled trivially (Figure 4). Additional ambiguous density was observed in the binding pocket of chain B, and is suspected to be low occupancy DAT180. Some key residue interactions of DAT180 are H-bonding with R221, pi-stacking with Y259 and H-bonding with Y261 in the binding pocket (Figure 5). There is unoccupied space in the pocket that can be exploited by adding chemical substituents to the

carboxyl tail of DAT180. This will have to be explored further with computational techniques such as free energy perturbation.

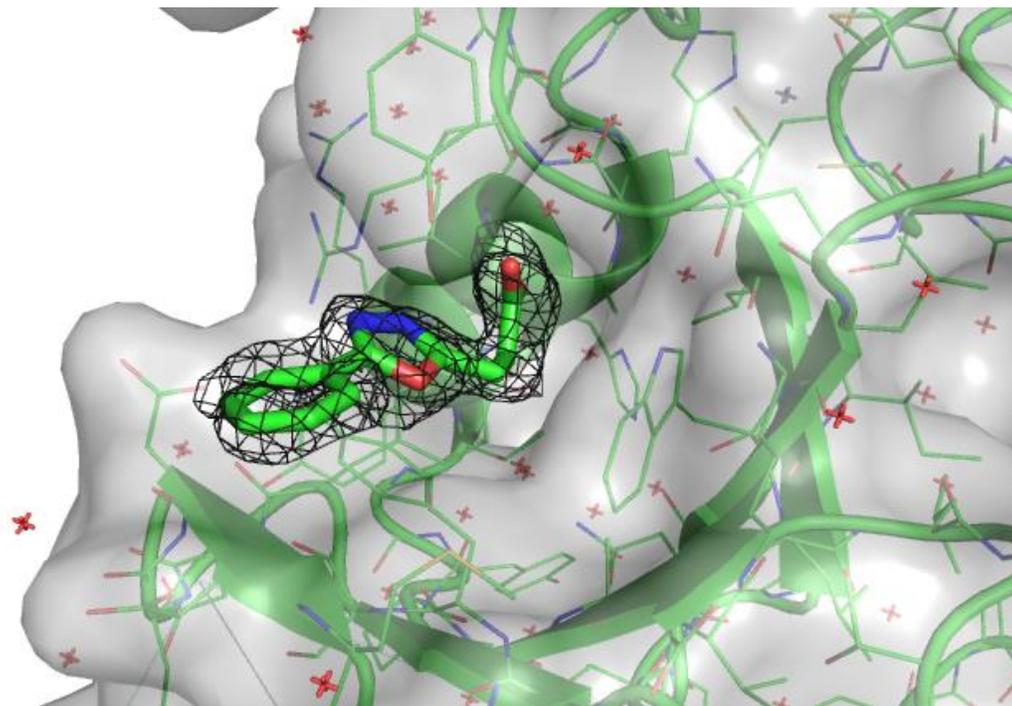


Figure 4: Omit map (σ_2) of USP5¹⁷¹⁻²⁹⁰ and DAT180

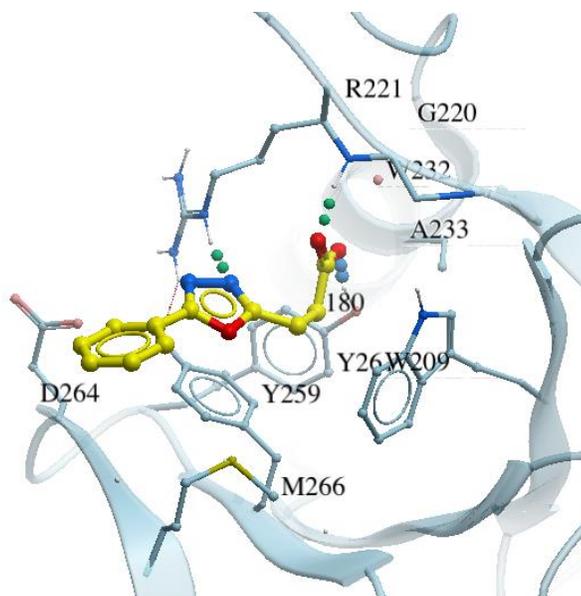


Figure 5. Key residue interactions of DAT180 with USP5¹⁷¹⁻²⁹⁰

DAT194

The solved crystal structure of DAT194 was the most interesting of the solved structures. We saw electron density in the binding pocket (chain A and B) but it did not correlate with the structure of the

compound we had ordered! After some discussion, we tried to see if modeling in an isomer of that compound would “fit” (Figure 6b). Both DAT194 and the Figure 6b isomer have the same mass. The purity of DAT194 was >95% when checked with UPLC/MS. It’s very cool to see that the high resolution crystal structure can differentiate between the electron densities of compounds to discern the correct structure of the compound. The omit map of the DAT194 isomer is shown in Figure 7.

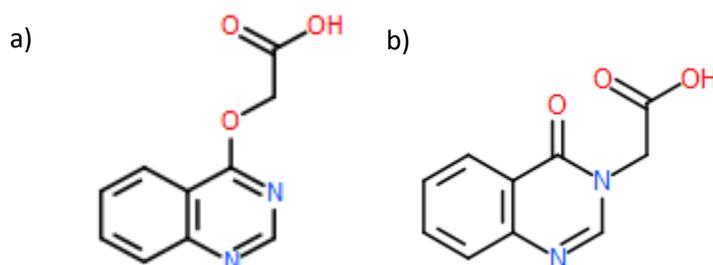


Figure 6. a) Expected DAT194 structure b) Modeled isomer with same mass as DAT194

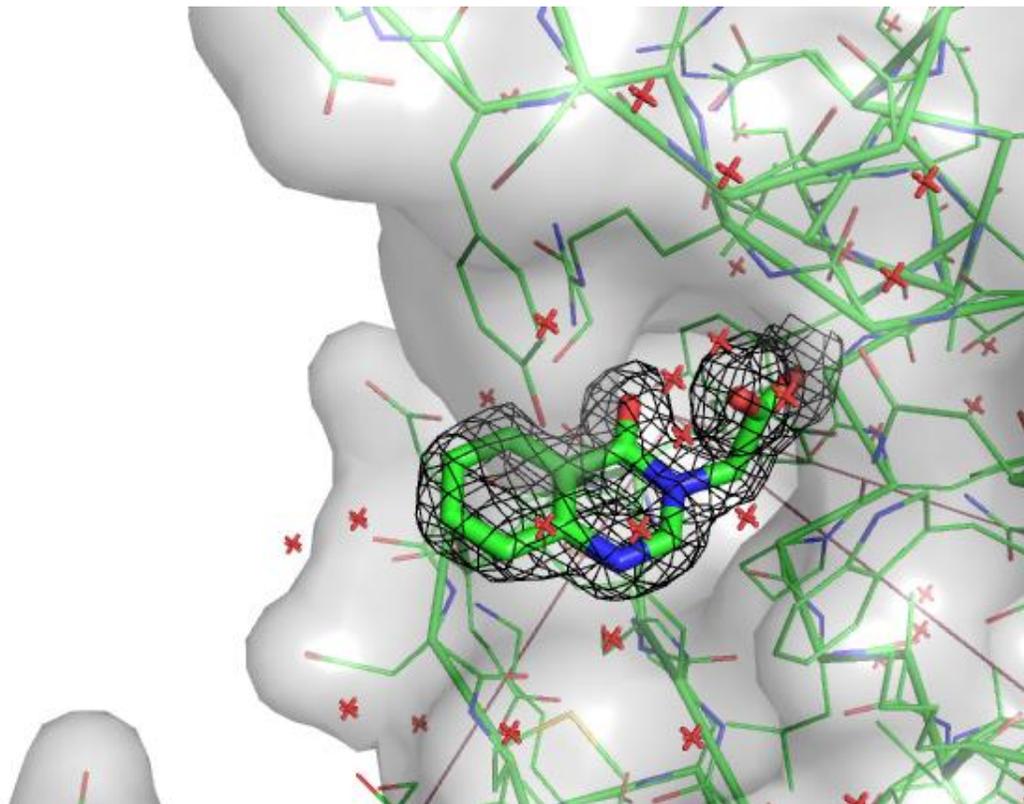


Figure 7. Omit map (σ_2) of USP5¹⁷¹⁻²⁹⁰ and isomer of DAT194

DAT194 H-bonds with R221 and has pi-stacking interactions with Y259 as well as coordination with W209, and Y261 (Figure 8).

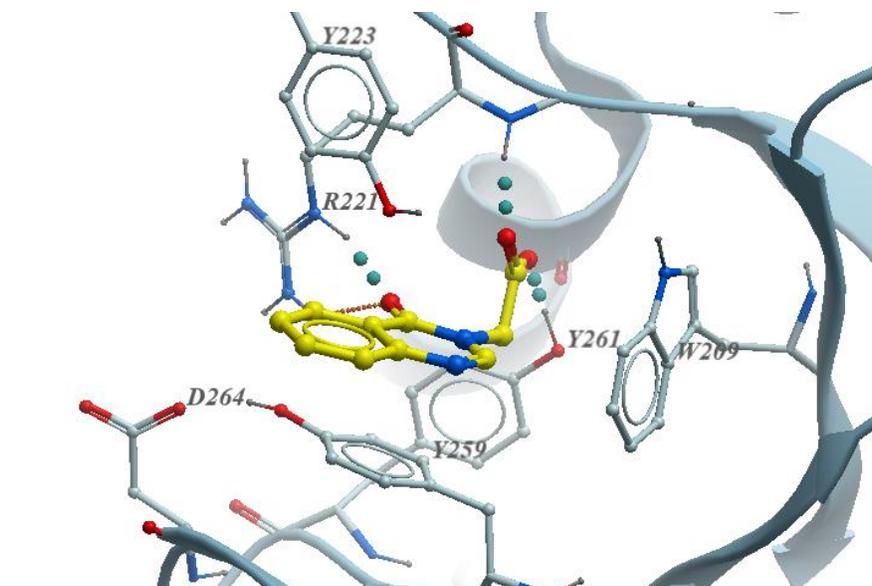


Figure 8. Key residue interactions of DAT194 with USP5¹⁷¹⁻²⁹⁰

DAT201

Electron density of DAT201 was seen clearly in the binding pocket of USP5¹⁷¹⁻²⁹⁰ (chain A) (Figure 9). Interestingly, DAT201 had a different crystal packing than the apo USP5 (PDB: [2G43](#)) and co-crystal structures of DAT180 and DAT194. A loop in the vicinity of the ligand was shifted in comparison to previous structures of USP5. The structure overlay of DAT201 with PDB [2G43](#) is shown in Figure 10. The loop shift in the co-crystal structure of DAT201 indicates the loop region is floppy in solution. It may be possible to exploit the area around the loop for design of follow-up compounds.

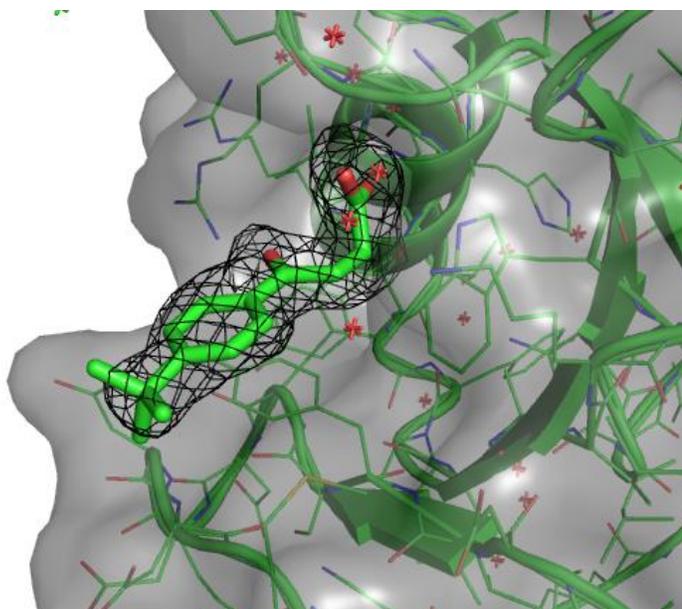


Figure 9. Omit map (σ_2) of USP5¹⁷¹⁻²⁹⁰ and isomer of DAT194

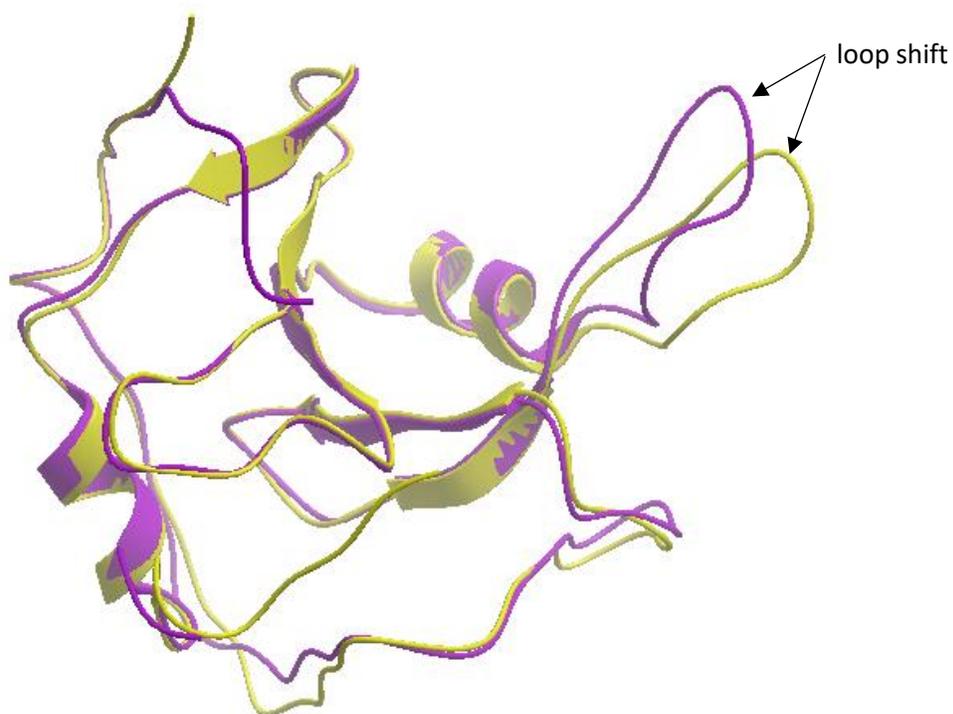


Figure 10. Overlay of DAT201 structure (yellow) & PDB: 2G43 (purple)

DAT201 H-bonds with R221 and pi-stacking with Y259 (Figure 11).

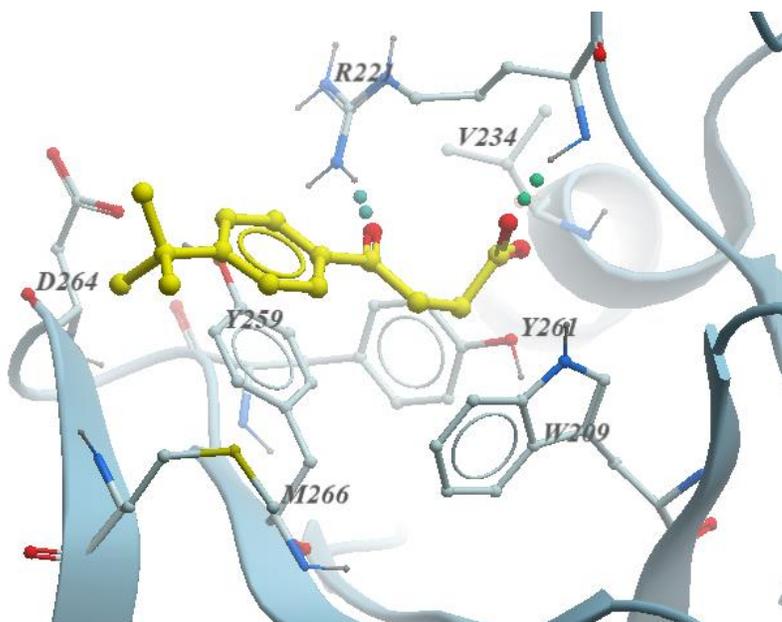


Figure 11. Key residue interactions of DAT201 with USP5¹⁷¹⁻²⁹⁰

Figure 12 shows an overlay of all three co-crystal structures to show the conserved features of ligand binding. All three ligands, DAT180, 194 and 201 have a conserved carboxylate tail that H-bonds with the backbone nitrogen of R221, and binds similarly to the C-terminal di-glycine motif of ubiquitin. All compounds are also engaged in a hydrogen-bond with the side-chain of R221, as does the di-glycine motif of ubiquitin indicating this is an important interaction in ligand design.

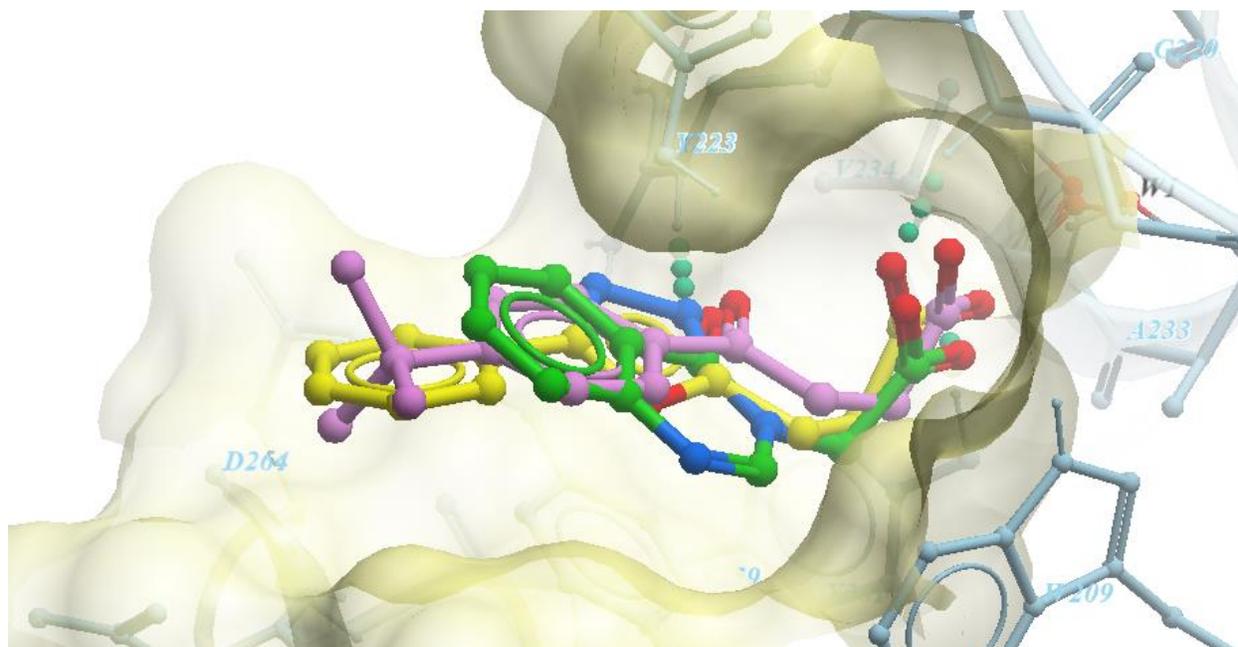


Figure 12. Overlay of co-crystal structures of DAT180 (yellow), 194 (green), 201 (purple)

Conclusions & Future Directions:

The solved crystal structures all had electron densities of compounds in the binding pocket; however, the structure of DAT194 was not as expected, but actually that of an isomer. In the future, I'll be using computational techniques to prioritize chemical analogs that may better exploit the binding pocket. Decorating the core scaffold of the compounds with different chemical groups can potentially increase ligand affinity to the protein.