PL-2016 Challenge Overview

Zied Gaieb

27 March, 2017

Topics

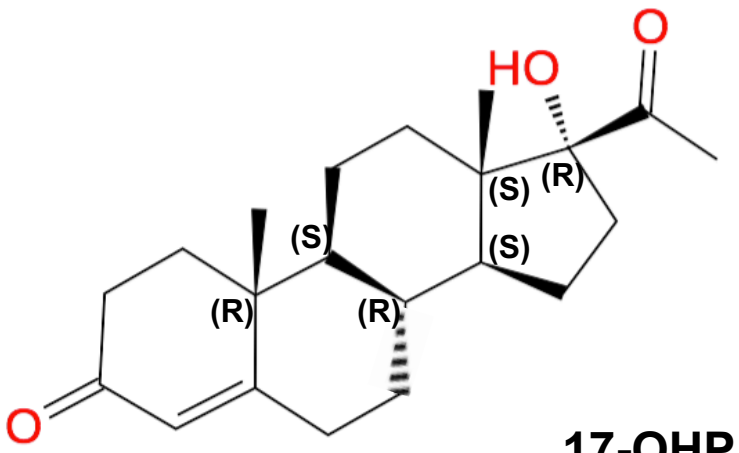
- PL-2016 Challenge Dataset Overview
- PL-2016 Challenge Instructions
- PL-2016 Challenge Results
- Conclusions

PL-2016 Challenge Dataset Overview

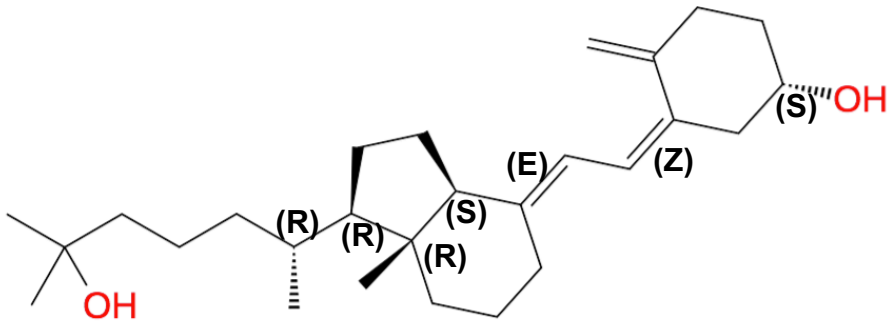
Challenge Dataset

- Kindly Donated by - Barry Stoddard (Fred Hutchinson Cancer Research) & David Baker (U. of Washington).

- Five protein structures designed using the Rosetta program.
  - **2** co-crystalized with the ligand 17-hydroxyprogesterone (17-OHP)
  - **3** co-crystalized with the ligand 25-hydroxycholecalciferol (25-D3).



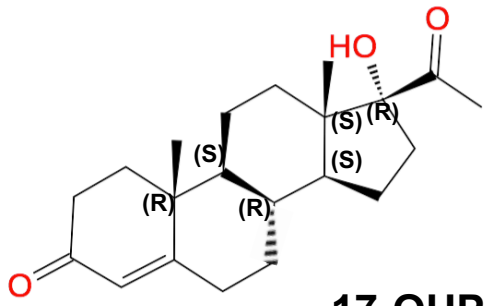
17-OHP



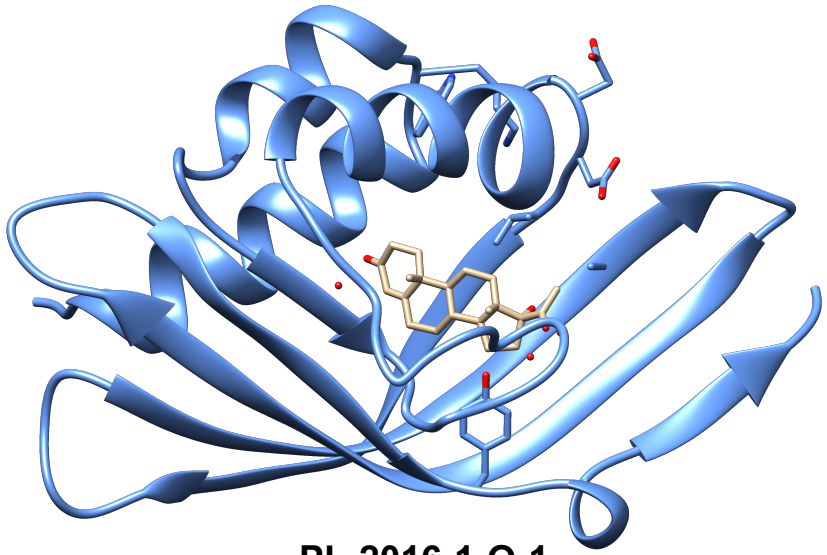
25-D3

The 17-OHP Binding Dataset

- In both structures:
  - The ligand is the same, 17-OHP
  - It is the designed proteins that differ, 8 mutations



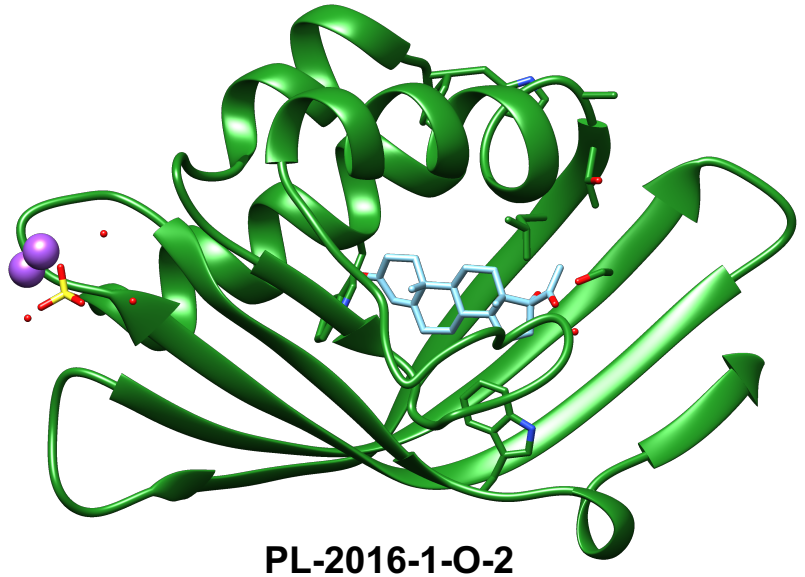
17-OHP



PL-2016-1-O-1

K_d = 60 ± 8 nM

Resolution = 2.5 Å; pH = 7.5



PL-2016-1-O-2

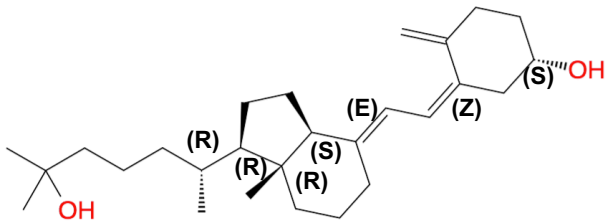
K_d = 15 ± 2 µM

Resolution = 2.0 Å; pH = 4.5

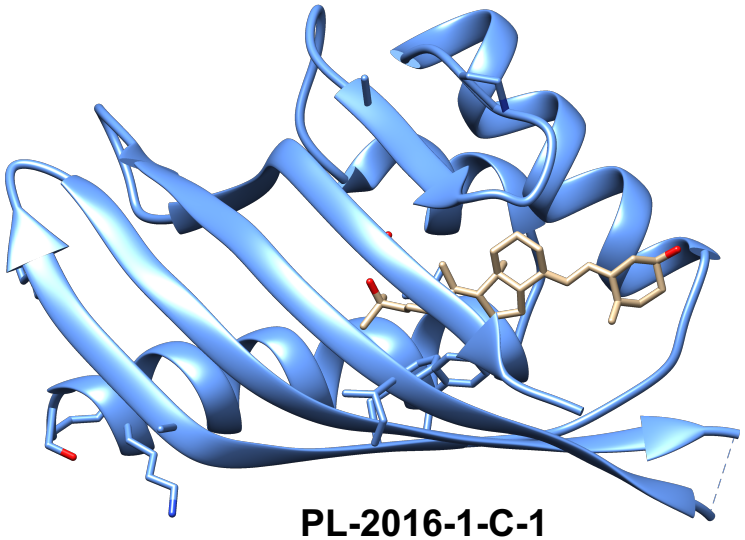
- 250 fold difference in activity.
- RMSD of ligands = 0.74 Å
- RMSD of proteins = 0.68 Å
- % Sequence identity = 93.13%

The cholecalciferol (25-D3) Binding Dataset

- In all structures:
  - The ligand is the same, 25-D3
  - It is the designed proteins that differ, 13 mutations

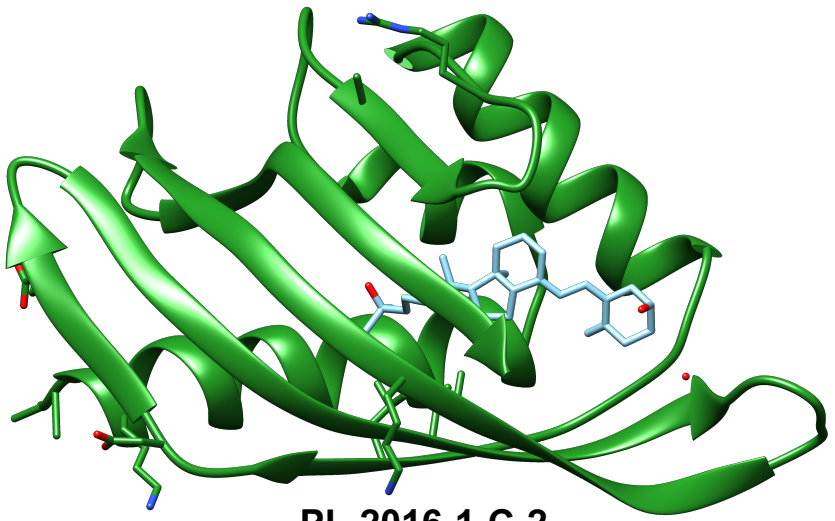


25-D3



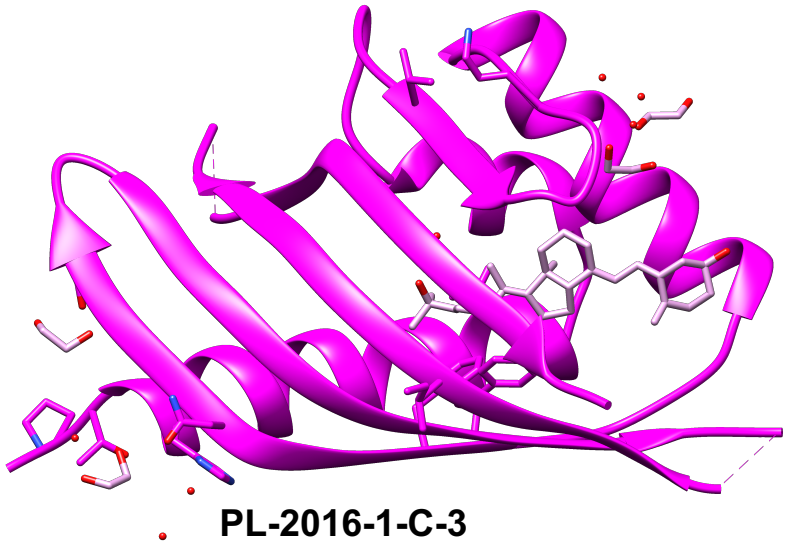
PL-2016-1-C-1

Resolution = 1.9 Å; pH = 4.6



PL-2016-1-C-2

Resolution = 2.1 Å; pH = 7.5



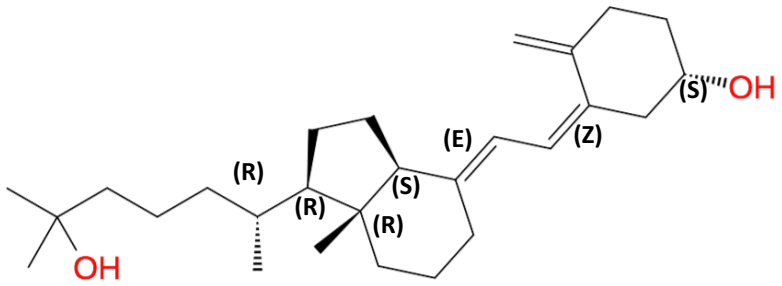
PL-2016-1-C-3

Resolution = 1.9 Å; pH = 7.5

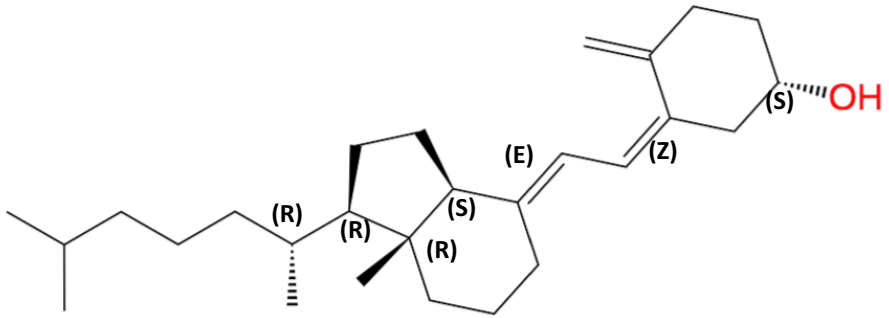
Protein 1	Protein 2	% identity	Protein RMSD (Å)	Ligand RMSD (Å)
PL-2016-1-C-1	PL-2016-1-C-2	92. 6	0.66	1.49
PL-2016-1-C-1	PL-2016-1-C-3	94.8	0.34	0.42
PL-2016-1-C-2	PL-2016-1-C-3	91.4	0.70	1.51

The cholecalciferol (25-D3) Binding Affinities

Protein	Ligand	Affinity (K_d)
PL-2016-1-C-1	25-D3	300 ± 40nm
PL-2016-1-C-2	25-D3	Similar to PL-2016-1-C-1
PL-2016-1-C-3	25-D3	Similar to PL-2016-1-C-1
PL-2016-1-C-1	**D3**	**~2 µM**



25-D3



D3

(No Crystal Structure)

PL-2016 Challenge Instructions

Challenge

- **Provided Inputs**
  - A) Protein structures with the co-crystalized ligand deleted and all the waters retained.
  - B) SDfile for ligands 17-OHP, 25-D3 and D3.

- **Outputs**
  - A) Predicted poses for the following complexes:
    - PL-2016-1-**O-1**/17-OHP
    - PL-2016-1-**O-2**/17-OHP
    - PL-2016-1-**C-1**/25-D3
    - PL-2016-1-**C-2**/25-D3
    - PL-2016-1-**C-3**/25-D3
  - B) Your predicted affinities, scores, or affinity rankings for the protein-ligand pairs for each of the two datasets

    Plus the affinity for PL-2016-1-C-1/vitamin D3.

What makes this challenge interesting?

- The OHP dataset has 250 fold difference in affinity even though the poses are very similar except for a few mutations and fewer crystal water in the binding site.

- The 25-D3 Dataset will also be compared in ranking to the vitamin D3 (6 fold difference in affinity)

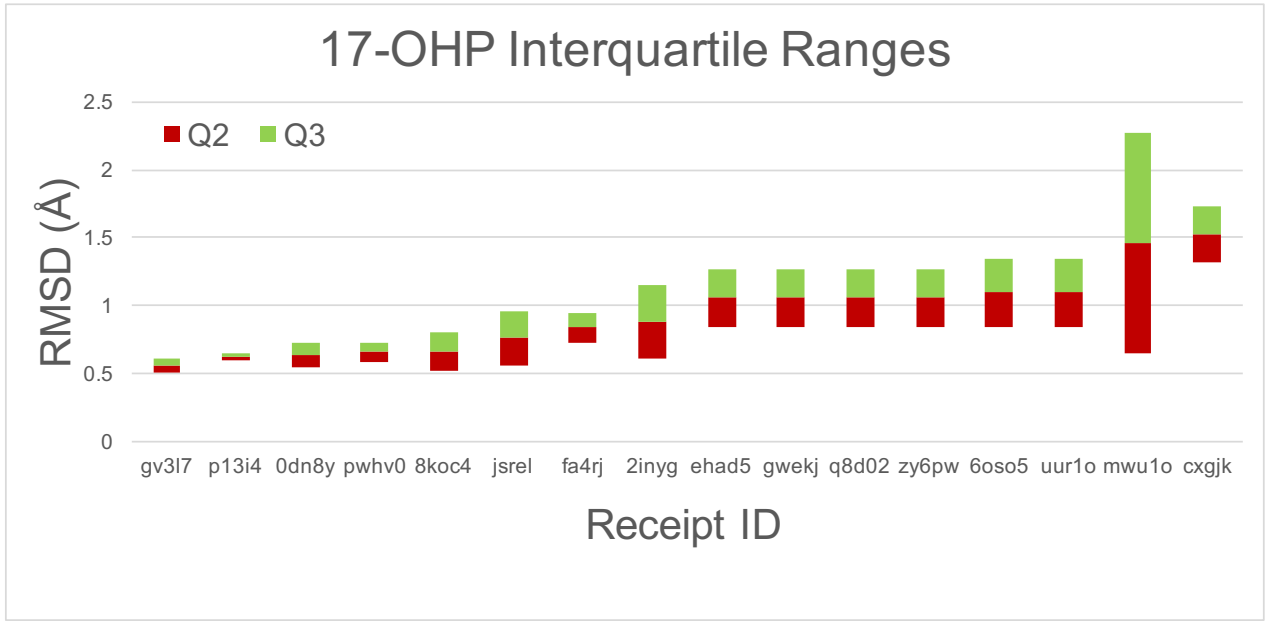
PL-2016 Challenge Results

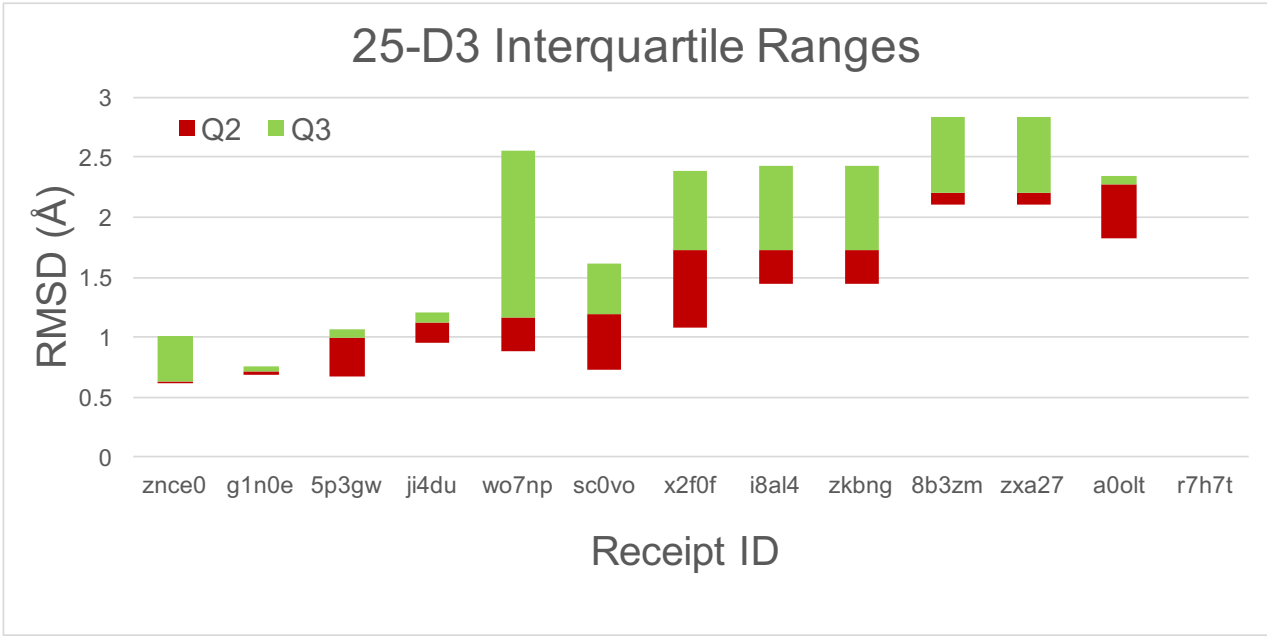
Methods Used by Participants

- A wide range of methods was used:
  - Docking methods: Glide-SP, Vina, Gold, Smina, MedusaDock, PIPER
  - Scoring methods: Glide, Vina, Gold score, MMGBSA, MMPBSA, many knowledge-based scoring methods

Pose Prediction Set Evaluation per Participant

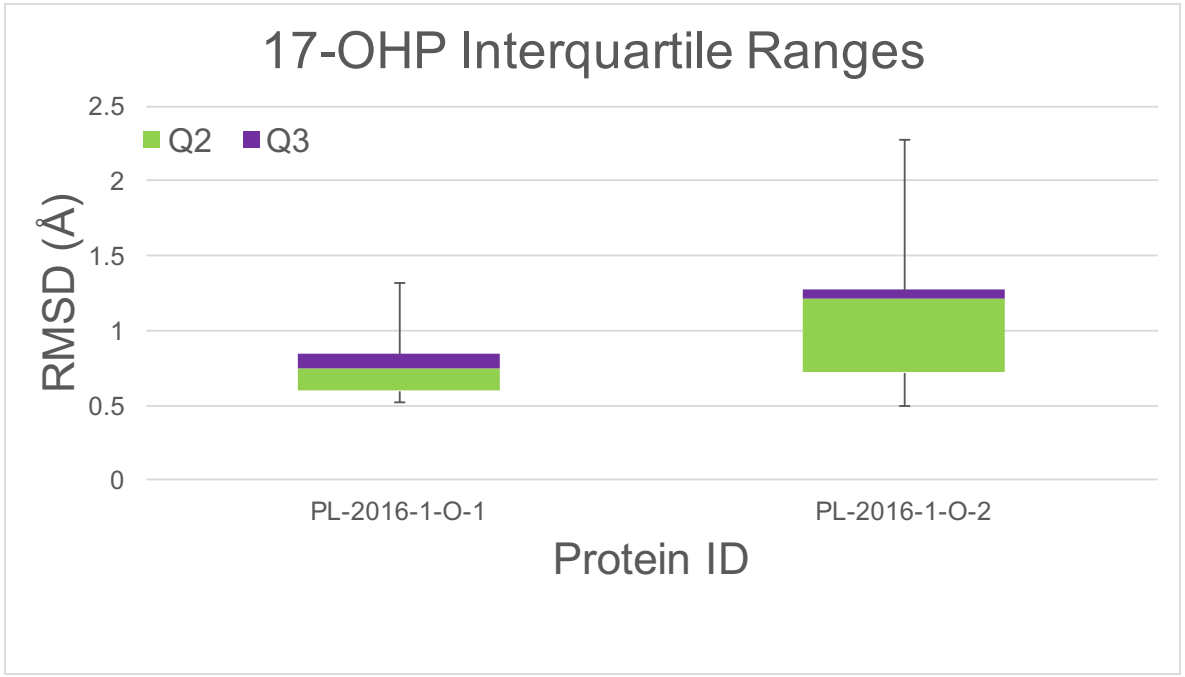
Ligand ID	Number of ligands	Number of participants
17-OHP	2	16
25-D3	3	13

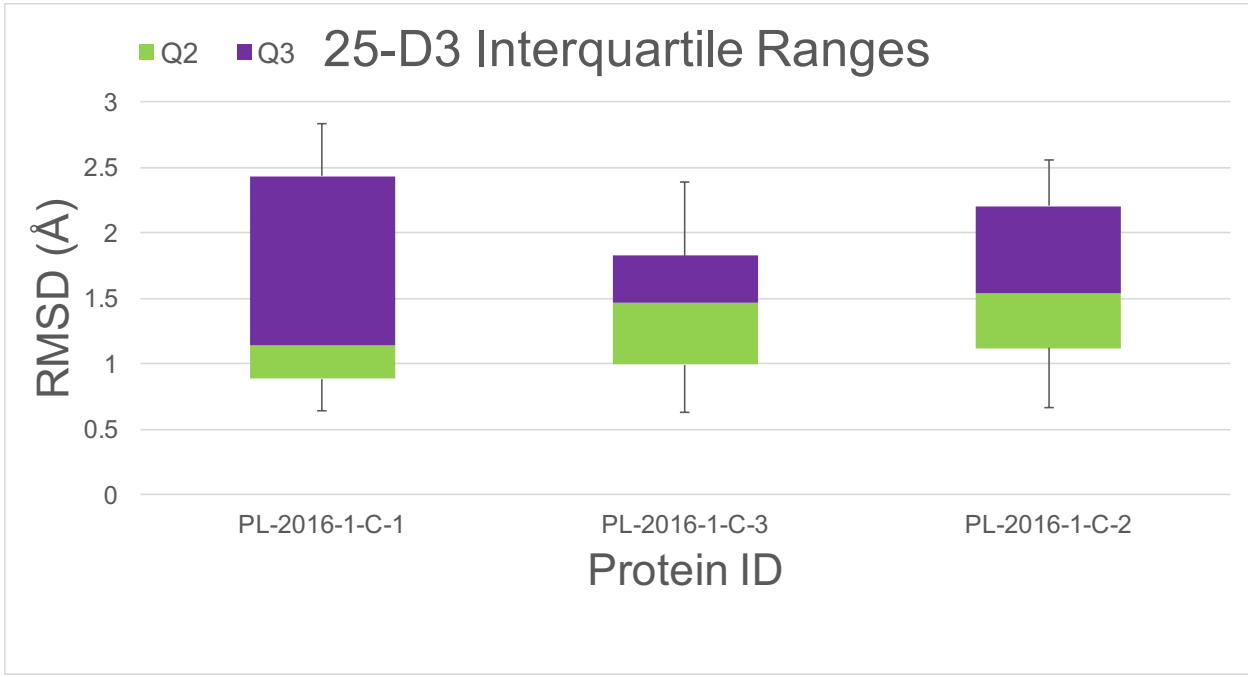




Pose Prediction Set Evaluation per Protein ID

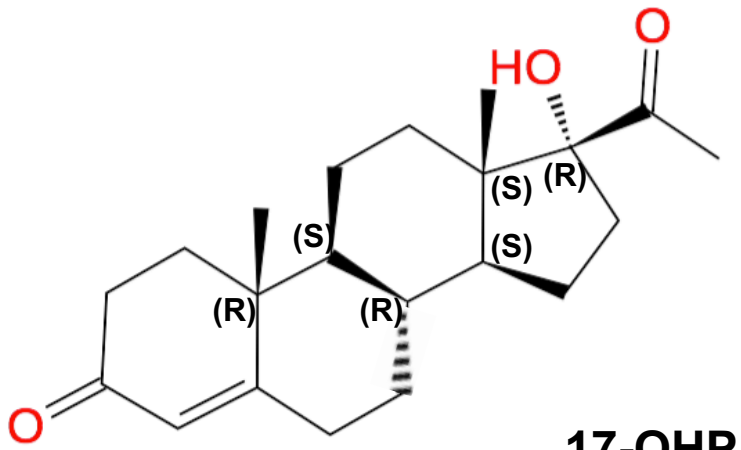
Ligand ID	Number of ligands	Number of participants
17-OHP	2	16
25-D3	3	13



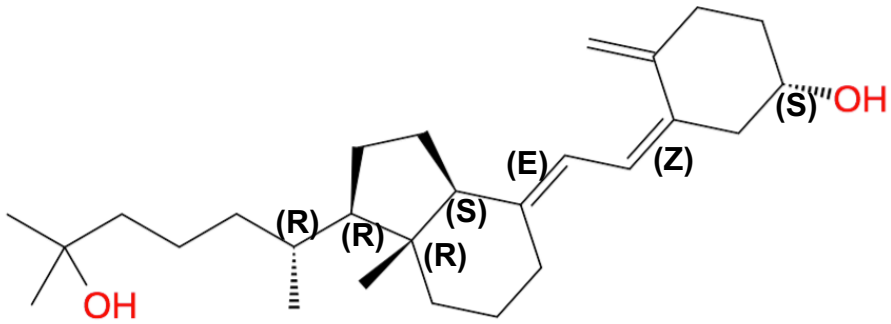


Scoring Set Evaluation

Ligand ID	Number of ligands scored	Number of participants	#with Kendall's Tau = 1
17-OHP	2	17	13/17
25-OHP	2	15	9/15



17-OHP



25-D3

Protein	Ligand	Affinity (K_d)
PL-2016-1-O-1	**17-OHP**	**60 ± 8 nM**
PL-2016-1-O-2	**17-OHP**	**15 ± 2 µM**

Protein	Ligand	Affinity (K_d)
PL-2016-1-C-1	**25-D3**	**300 ± 40nm**
PL-2016-1-C-2	25-D3	Similar to PL-2016-1-C-1
PL-2016-1-C-3	25-D3	Similar to PL-2016-1-C-1
PL-2016-1-C-1	**D3**	**~2 µM**

Conclusions

What we learnt

- Pose prediction – Average of the Mean RMSD of Best pose was:
  - 0.9 Å for 17-OHP binders
  - 1.5 Å for the 25-D3 binders

- Scoring prediction – More than half of the predictions ranked the ligand scores correctly in both targets.

- This was a self-docking challenge and that would explain the good performance across the various methods used.