Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 11:09 PM EDT

PDB ID	:	6U9S
Title	:	Crystal structure of human CD81 large extracellular loop in complex with 5A6 Fab
Authors	:	Susa, K.J.; Seegar, T.C.M.; Blacklow, S.C.B.; Kruse, A.C.
Deposited on	:	2019-09-09
Resolution	:	2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8374 atoms, of which 24 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called 5A6 FAB Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	221	1641	1044	265	327	5	0	0	0
1	D	218	1593	1012	255	321	5	0	0	0

- Molecule 2 is a protein called 5A6 FAB Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	219	1674	1051	279	338	6	0	0	0
2	E	218	1666	1045	279	337	5	0	0	0

- Molecule 3 is a protein called CD81 antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	92	701	433	119	145	4	0	0	0
3	F	92	699	432	119	144	4	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

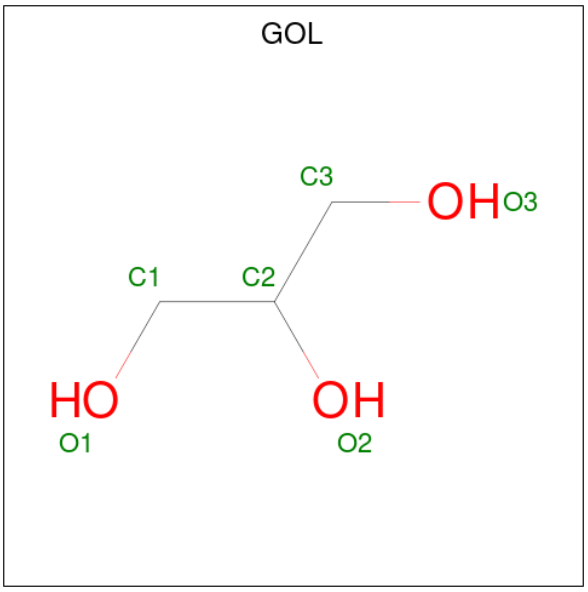
Chain	Residue	Modelled	Actual	Comment	Reference
C	110	GLU	-	expression tag	UNP P60033
C	111	THR	-	expression tag	UNP P60033
C	201	SER	-	expression tag	UNP P60033
C	202	HIS	-	expression tag	UNP P60033
C	203	HIS	-	expression tag	UNP P60033
C	204	HIS	-	expression tag	UNP P60033
C	205	HIS	-	expression tag	UNP P60033
C	206	HIS	-	expression tag	UNP P60033

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	207	HIS	-	expression tag	UNP P60033
F	110	GLU	-	expression tag	UNP P60033
F	111	THR	-	expression tag	UNP P60033
F	201	SER	-	expression tag	UNP P60033
F	202	HIS	-	expression tag	UNP P60033
F	203	HIS	-	expression tag	UNP P60033
F	204	HIS	-	expression tag	UNP P60033
F	205	HIS	-	expression tag	UNP P60033
F	206	HIS	-	expression tag	UNP P60033
F	207	HIS	-	expression tag	UNP P60033

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	14	3	8	3	0	0
			Total	C	H	O		
4	B	1	14	3	8	3	0	0
			Total	C	H	O		
4	E	1	14	3	8	3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	77	77	77	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	B	94	94	94	0	0
			Total	O		
5	C	69	69	69	0	0
			Total	O		
5	D	25	25	25	0	0
			Total	O		
5	E	40	40	40	0	0
			Total	O		
5	F	53	53	53	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.00Å 96.86Å 297.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.52 – 2.40	Depositor
% Data completeness (in resolution range)	87.0 (49.52-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.67 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.16_3549, PHENIX 1.16_3549	Depositor
R, R_{free}	0.214 , 0.281	Depositor
Wilson B-factor (Å^2)	42.7	Xtriage
Anisotropy	0.256	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8374	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	GOL	A	401	-	5,5,5	0.71	0	5,5,5	0.85	0
4	GOL	B	301	-	5,5,5	0.72	0	5,5,5	0.88	0
4	GOL	E	301	-	5,5,5	0.78	0	5,5,5	0.95	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	401	-	-	2/4/4/4	-
4	GOL	B	301	-	-	4/4/4/4	-
4	GOL	E	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	GOL	C1-C2-C3-O3
4	B	301	GOL	O1-C1-C2-C3
4	B	301	GOL	C1-C2-C3-O3
4	E	301	GOL	O1-C1-C2-C3
4	B	301	GOL	O2-C2-C3-O3
4	A	401	GOL	O2-C2-C3-O3
4	B	301	GOL	O1-C1-C2-O2
4	E	301	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates (i)

EDS failed to run properly - this section is therefore empty.

5.4 Ligands (i)

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers (i)

EDS failed to run properly - this section is therefore empty.