Full wwPDB X-ray Structure Validation Report (i)

Oct 3, 2023 – 02:58 AM EDT

PDB ID	:	6UH4
Title	:	B. theta Bile Salt Hydrolase with covalent inhibitor
Authors	:	Seegar, T.C.M.
Deposited on	:	2019-09-26
Resolution	:	3.51 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 (1)) were used in the production of this report:

MolProbity	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
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 3.51 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 2 unique types of molecules in this entry. The entry contains 10290 atoms, of which 41 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 Choloylglycine hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	0	0	0
			2597	1653	437	491	16			
1	B	326	Total	C	N	O	S	0	0	0
			2574	1637	434	488	15			
1	C	319	Total	C	N	O	S	0	0	0
			2525	1608	423	480	14			
1	D	319	Total	C	N	O	S	0	0	0
			2525	1608	423	480	14			

There are 36 discrepancies between the modelled and reference sequences:

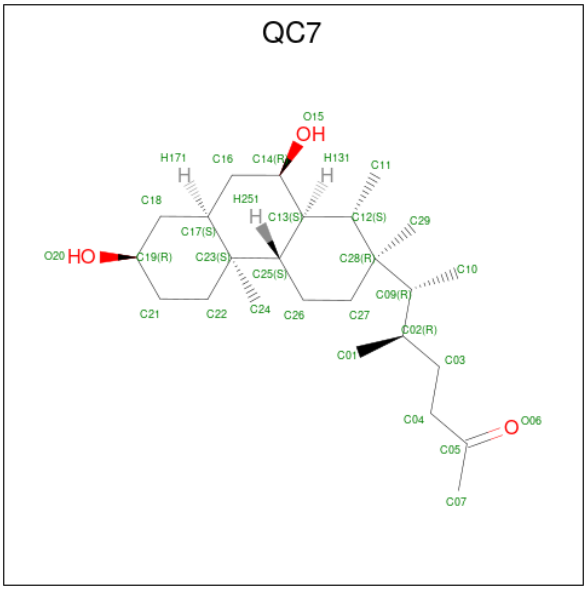
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A0P0ENF5
A	329	LEU	-	expression tag	UNP A0A0P0ENF5
A	330	GLU	-	expression tag	UNP A0A0P0ENF5
A	331	HIS	-	expression tag	UNP A0A0P0ENF5
A	332	HIS	-	expression tag	UNP A0A0P0ENF5
A	333	HIS	-	expression tag	UNP A0A0P0ENF5
A	334	HIS	-	expression tag	UNP A0A0P0ENF5
A	335	HIS	-	expression tag	UNP A0A0P0ENF5
A	336	HIS	-	expression tag	UNP A0A0P0ENF5
B	1	MET	-	initiating methionine	UNP A0A0P0ENF5
B	329	LEU	-	expression tag	UNP A0A0P0ENF5
B	330	GLU	-	expression tag	UNP A0A0P0ENF5
B	331	HIS	-	expression tag	UNP A0A0P0ENF5
B	332	HIS	-	expression tag	UNP A0A0P0ENF5
B	333	HIS	-	expression tag	UNP A0A0P0ENF5
B	334	HIS	-	expression tag	UNP A0A0P0ENF5
B	335	HIS	-	expression tag	UNP A0A0P0ENF5
B	336	HIS	-	expression tag	UNP A0A0P0ENF5
C	1	MET	-	initiating methionine	UNP A0A0P0ENF5
C	329	LEU	-	expression tag	UNP A0A0P0ENF5
C	330	GLU	-	expression tag	UNP A0A0P0ENF5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	331	HIS	-	expression tag	UNP A0A0P0ENF5
C	332	HIS	-	expression tag	UNP A0A0P0ENF5
C	333	HIS	-	expression tag	UNP A0A0P0ENF5
C	334	HIS	-	expression tag	UNP A0A0P0ENF5
C	335	HIS	-	expression tag	UNP A0A0P0ENF5
C	336	HIS	-	expression tag	UNP A0A0P0ENF5
D	1	MET	-	initiating methionine	UNP A0A0P0ENF5
D	329	LEU	-	expression tag	UNP A0A0P0ENF5
D	330	GLU	-	expression tag	UNP A0A0P0ENF5
D	331	HIS	-	expression tag	UNP A0A0P0ENF5
D	332	HIS	-	expression tag	UNP A0A0P0ENF5
D	333	HIS	-	expression tag	UNP A0A0P0ENF5
D	334	HIS	-	expression tag	UNP A0A0P0ENF5
D	335	HIS	-	expression tag	UNP A0A0P0ENF5
D	336	HIS	-	expression tag	UNP A0A0P0ENF5

- Molecule 2 is (5R,6R)-6-[(1S,2R,4aS,4bS,7R,8aS,10R,10aS)-7,10-dihydroxy-1,2,4b-trimethyl tetradecahydrophenanthren-2-yl]-5-methylheptan-2-one (three-letter code: QC7) (formula: $C_{25}H_{44}O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	B	1	69	25	41	3	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 2	Depositor
Cell constants a, b, c, α, β, γ	98.58Å 99.52Å 162.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.29 – 3.51	Depositor
% Data completeness (in resolution range)	97.0 (49.29-3.51)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.238 , 0.292	Depositor
Wilson B-factor (Å^2)	142.4	Xtriage
Anisotropy	0.563	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.35$, $< L^2 > = 0.18$	Xtriage
Estimated twinning fraction	0.226 for k,h,-l	Xtriage
Total number of atoms	10290	wwPDB-VP
Average B, all atoms (Å^2)	187.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.91% 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

4.1 Standard geometry

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

4.2 Too-close contacts

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

4.3 Torsion angles

4.3.1 Protein backbone

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

4.3.2 Protein sidechains

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

4.3.3 RNA

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

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no monosaccharides in this entry.

4.6 Ligand geometry

1 ligand is 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	$\#\|Z\| > 2$
2	QC7	B	401	1	28,30,30	2.77	8 (28%)	40,47,47	2.88	15 (37%)

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
2	QC7	B	401	1	-	4/15/65/65	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	QC7	C27-C26	7.74	1.69	1.53
2	B	401	QC7	C13-C25	6.69	1.66	1.53
2	B	401	QC7	C16-C14	5.48	1.62	1.52
2	B	401	QC7	C16-C17	4.85	1.61	1.53
2	B	401	QC7	C11-C12	3.95	1.61	1.53
2	B	401	QC7	C13-C14	-2.35	1.49	1.53
2	B	401	QC7	C10-C09	2.31	1.58	1.53
2	B	401	QC7	O15-C14	-2.18	1.38	1.43

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	QC7	C11-C12-C28	-9.87	102.96	114.62
2	B	401	QC7	C11-C12-C13	6.34	121.22	111.87
2	B	401	QC7	C18-C17-C16	-6.17	104.08	111.19
2	B	401	QC7	C28-C12-C13	5.77	118.03	111.81
2	B	401	QC7	C25-C13-C12	-4.64	103.32	109.72
2	B	401	QC7	C03-C04-C05	-4.35	110.09	114.57
2	B	401	QC7	C25-C13-C14	3.53	116.10	111.88
2	B	401	QC7	C04-C03-C02	2.84	119.70	114.52
2	B	401	QC7	C26-C27-C28	2.62	117.27	112.78
2	B	401	QC7	C23-C25-C13	2.59	114.60	111.82
2	B	401	QC7	C22-C23-C17	2.51	111.47	107.77
2	B	401	QC7	C17-C16-C14	-2.40	111.81	114.46
2	B	401	QC7	C22-C21-C19	2.11	113.18	110.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	QC7	C29-C28-C12	-2.05	107.59	110.94
2	B	401	QC7	C22-C23-C25	-2.02	108.18	111.35

There are no chirality outliers.

All (4) torsion outliers are listed below:

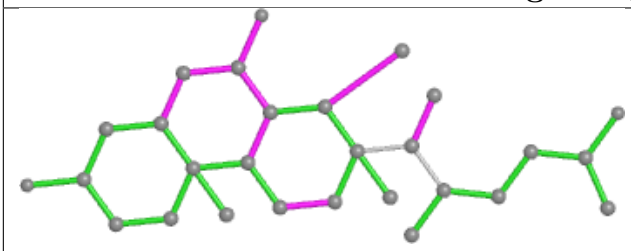
Mol	Chain	Res	Type	Atoms
2	B	401	QC7	C03-C04-C05-C07
2	B	401	QC7	C03-C04-C05-O06
2	B	401	QC7	C10-C09-C28-C29
2	B	401	QC7	C10-C09-C28-C12

There are no ring outliers.

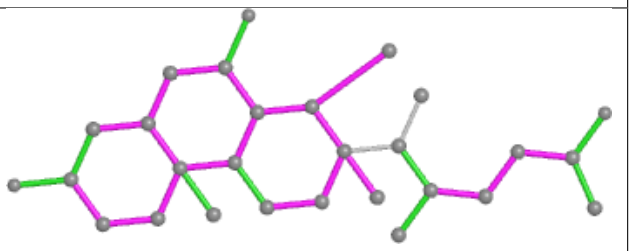
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

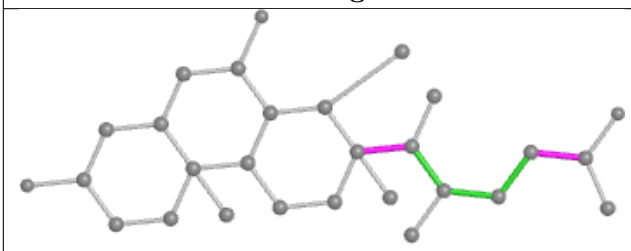
Ligand QC7 B 401



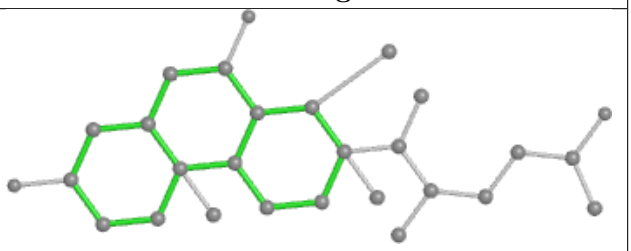
Bond lengths



Bond angles



Torsions



Rings

4.7 Other polymers

There are no such residues in this entry.

4.8 Polymer linkage issues

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

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

5.2 Non-standard residues in protein, DNA, RNA chains

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

5.3 Carbohydrates

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

5.4 Ligands

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

5.5 Other polymers

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