wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2023 – 04:15 AM EDT

PDB ID	:	6VUY
Title	:	Crystal structure of Eis from Mycobacterium tuberculosis in complex with inhibitor SGT358
Authors	:	Punetha, A.; Hou, C.; Ngo, H.X.; Garneau-Tsodikova, S.; Tsodikov, O.V.
Deposited on	:	2020-02-16
Resolution	:	2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 2.70 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 8 unique types of molecules in this entry. The entry contains 3300 atoms, of which 0 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.

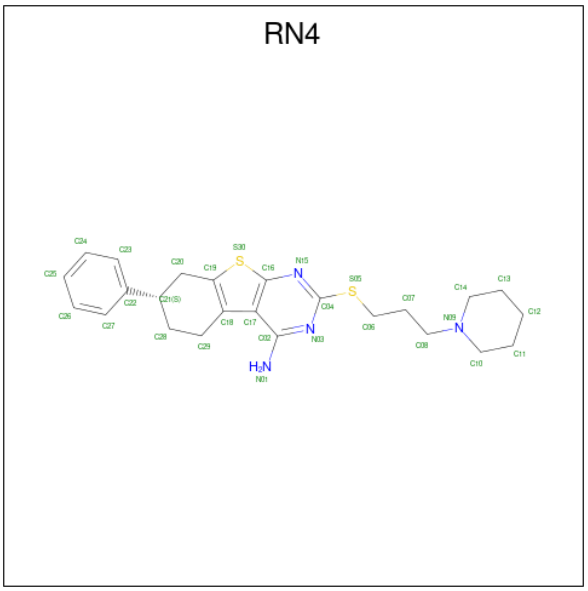
- Molecule 1 is a protein called N-acetyltransferase Eis.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	396	3051	1917	563	560	11	0	1	0

There are 21 discrepancies between the modelled and reference sequences:

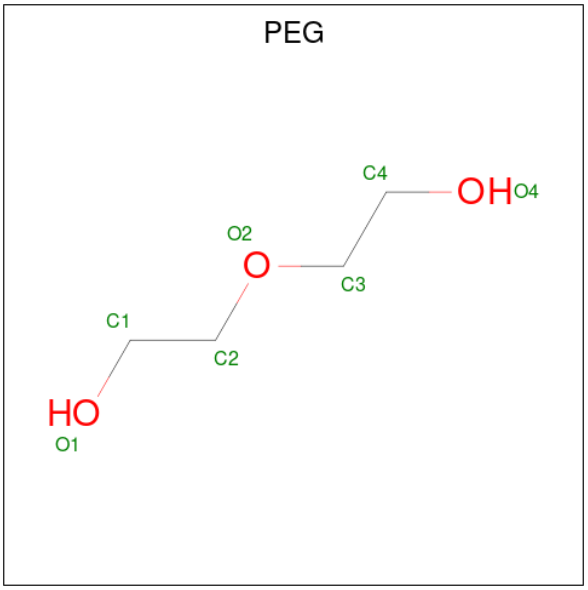
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP P9WFK7
A	-18	GLY	-	expression tag	UNP P9WFK7
A	-17	SER	-	expression tag	UNP P9WFK7
A	-16	SER	-	expression tag	UNP P9WFK7
A	-15	HIS	-	expression tag	UNP P9WFK7
A	-14	HIS	-	expression tag	UNP P9WFK7
A	-13	HIS	-	expression tag	UNP P9WFK7
A	-12	HIS	-	expression tag	UNP P9WFK7
A	-11	HIS	-	expression tag	UNP P9WFK7
A	-10	HIS	-	expression tag	UNP P9WFK7
A	-9	SER	-	expression tag	UNP P9WFK7
A	-8	SER	-	expression tag	UNP P9WFK7
A	-7	GLY	-	expression tag	UNP P9WFK7
A	-6	LEU	-	expression tag	UNP P9WFK7
A	-5	VAL	-	expression tag	UNP P9WFK7
A	-4	PRO	-	expression tag	UNP P9WFK7
A	-3	ARG	-	expression tag	UNP P9WFK7
A	-2	GLY	-	expression tag	UNP P9WFK7
A	-1	SER	-	expression tag	UNP P9WFK7
A	0	HIS	-	expression tag	UNP P9WFK7
A	204	ALA	CYS	engineered mutation	UNP P9WFK7

- Molecule 2 is (7S)-7-phenyl-2-{[3-(piperidin-1-yl)propyl]sulfanyl}-5,6,7,8-tetrahydro[1]benzothieno[2,3-d]pyrimidin-4-amine (three-letter code: RN4) (formula: $C_{24}H_{30}N_4S_2$) (labeled as "Ligand of Interest" by depositor).



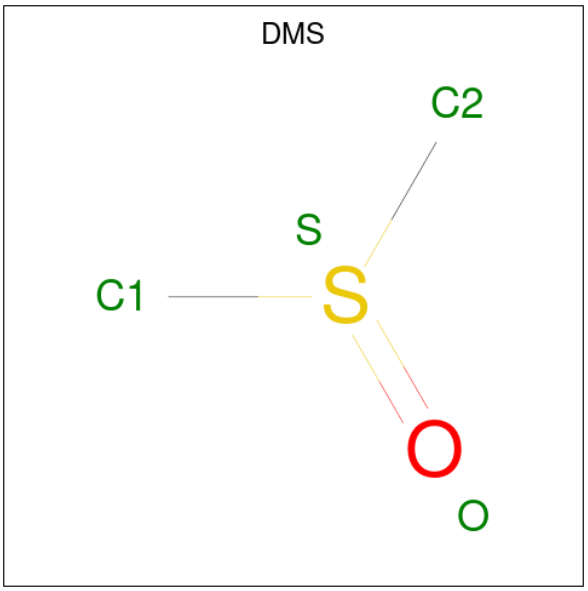
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	S		
2	A	1	30	24	4	2	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



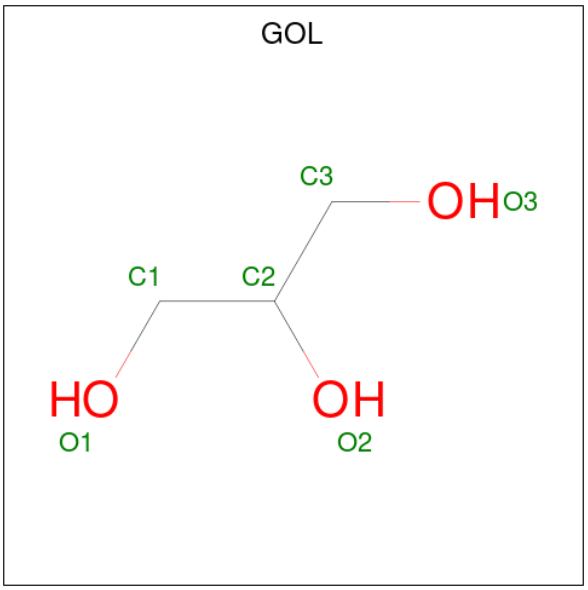
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
3	A	1	7	4	3		0	0
			Total	C	O			
3	A	1	7	4	3		0	0

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
4	A	1	4	2	1	1	0	0
			Total	C	O	S		
4	A	1	4	2	1	1	0	0
			Total	C	O	S		
4	A	1	4	2	1	1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	6	3	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	6	3	3	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
6	A	4	4	4	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
7	A	1	1	1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	176	176	176	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	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	175.05Å 175.05Å 123.56Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.29 – 2.70	Depositor
% Data completeness (in resolution range)	98.0 (37.29-2.70)	Depositor
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.171 , 0.213	Depositor
Wilson B-factor (Å^2)	38.2	Xtriage
Anisotropy	0.727	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	3300	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% 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)

Of 13 ligands modelled in this entry, 5 are monoatomic - leaving 8 for Mogul analysis.

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	$\#\|Z\| > 2$
5	GOL	A	507	-	5,5,5	0.15	0	5,5,5	0.37	0
4	DMS	A	506	-	3,3,3	0.32	0	3,3,3	0.04	0
4	DMS	A	504	-	3,3,3	0.31	0	3,3,3	0.11	0
5	GOL	A	508	-	5,5,5	0.13	0	5,5,5	0.35	0
3	PEG	A	502	-	6,6,6	0.29	0	5,5,5	0.19	0
3	PEG	A	503	-	6,6,6	0.29	0	5,5,5	0.19	0
4	DMS	A	505	-	3,3,3	0.26	0	3,3,3	0.04	0
2	RN4	A	501	-	32,34,34	2.44	5 (15%)	31,47,47	1.86	8 (25%)

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
5	GOL	A	507	-	-	2/4/4/4	-
5	GOL	A	508	-	-	2/4/4/4	-
3	PEG	A	502	-	-	2/4/4/4	-
3	PEG	A	503	-	-	2/4/4/4	-
2	RN4	A	501	-	-	4/11/28/28	0/5/5/5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	RN4	C08-N09	-10.34	1.23	1.47
2	A	501	RN4	C18-C17	4.87	1.53	1.40
2	A	501	RN4	C04-S05	3.52	1.82	1.75
2	A	501	RN4	C20-C19	3.33	1.53	1.50
2	A	501	RN4	C02-N01	3.10	1.45	1.34

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	RN4	C19-C20-C21	-4.99	108.56	113.71
2	A	501	RN4	C28-C21-C22	4.18	122.60	112.79
2	A	501	RN4	C11-C10-N09	3.41	116.80	111.33
2	A	501	RN4	C06-S05-C04	2.98	106.95	102.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	RN4	N15-C04-N03	-2.86	121.72	126.98

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

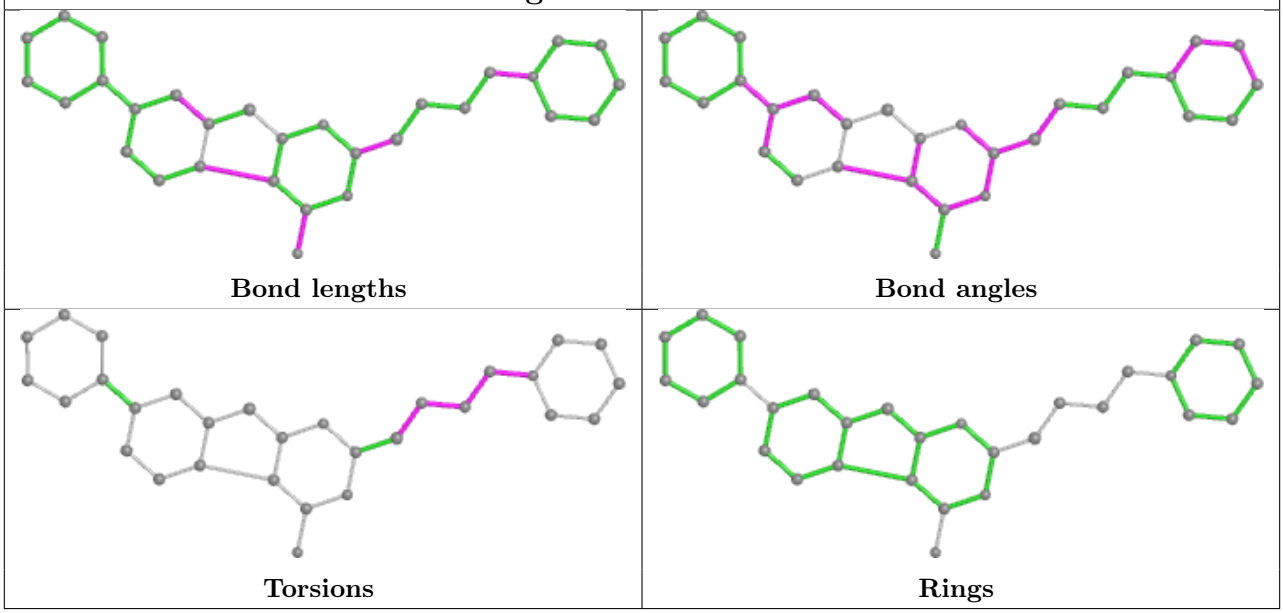
Mol	Chain	Res	Type	Atoms
2	A	501	RN4	S05-C06-C07-C08
2	A	501	RN4	C06-C07-C08-N09
5	A	507	GOL	C1-C2-C3-O3
5	A	508	GOL	O1-C1-C2-C3
3	A	502	PEG	O2-C3-C4-O4

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 RN4 A 501



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.