



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:16 AM GMT

PDB ID : 3OB7
Title : Human Thymidylate Synthase R163K with Cys 195 covalently modified by Glutathione
Authors : Gibson, L.M.; Celeste, L.R.; Lovelace, L.L.; Lebioda, L.
Deposited on : 2010-08-06
Resolution : 2.75 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

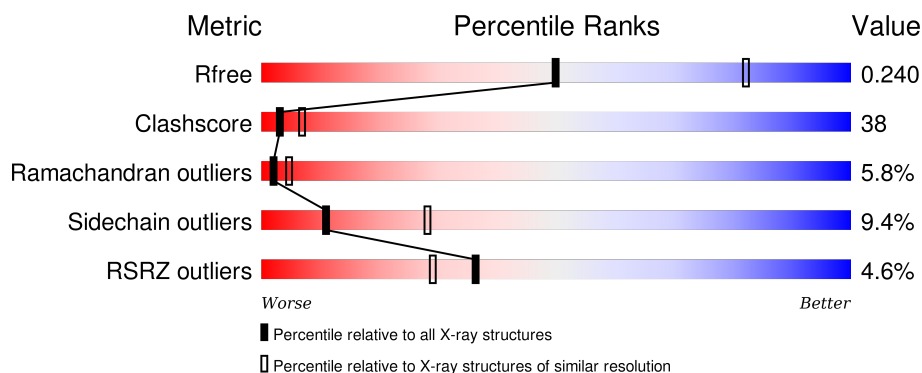
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3340 (2.80-2.72)
Clashscore	102246	3829 (2.80-2.72)
Ramachandran outliers	100387	3767 (2.80-2.72)
Sidechain outliers	100360	3770 (2.80-2.72)
RSRZ outliers	91569	3352 (2.80-2.72)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	313	<div> <div>50%</div> <div>37%</div> <div>10%</div> </div>
1	B	313	<div>4%</div> <div>40%</div> <div>41%</div> <div>9%</div> <div>10%</div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GSH	A	581	-	-	-	X
2	GSH	D	314	-	-	X	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11445 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.

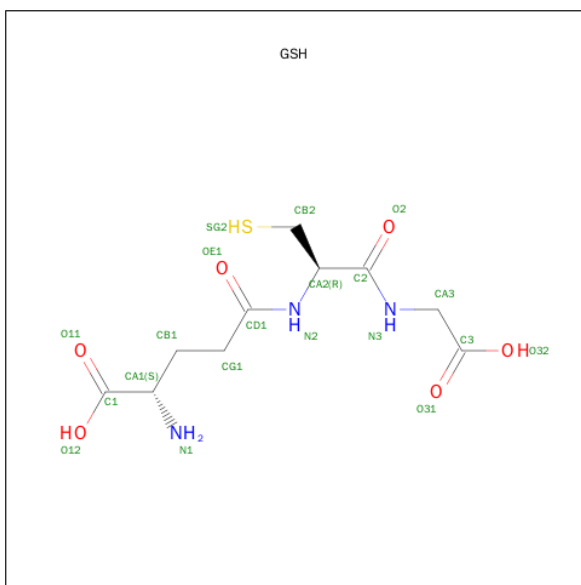
- Molecule 1 is a protein called Thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	281	Total	C	N	O	S	0	0	0
			2268	1451	395	411	11			
1	B	281	Total	C	N	O	S	0	0	0
			2268	1451	395	411	11			
1	C	280	Total	C	N	O	S	0	0	0
			2261	1447	394	409	11			
1	D	281	Total	C	N	O	S	0	0	0
			2268	1451	395	411	11			
1	E	280	Total	C	N	O	S	0	0	0
			2261	1447	394	409	11			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	163	LYS	ARG	ENGINEERED MUTATION	UNP P04818
B	163	LYS	ARG	ENGINEERED MUTATION	UNP P04818
C	163	LYS	ARG	ENGINEERED MUTATION	UNP P04818
D	163	LYS	ARG	ENGINEERED MUTATION	UNP P04818
E	163	LYS	ARG	ENGINEERED MUTATION	UNP P04818

- Molecule 2 is GLUTATHIONE (three-letter code: GSH) (formula: C₁₀H₁₇N₃O₆S).

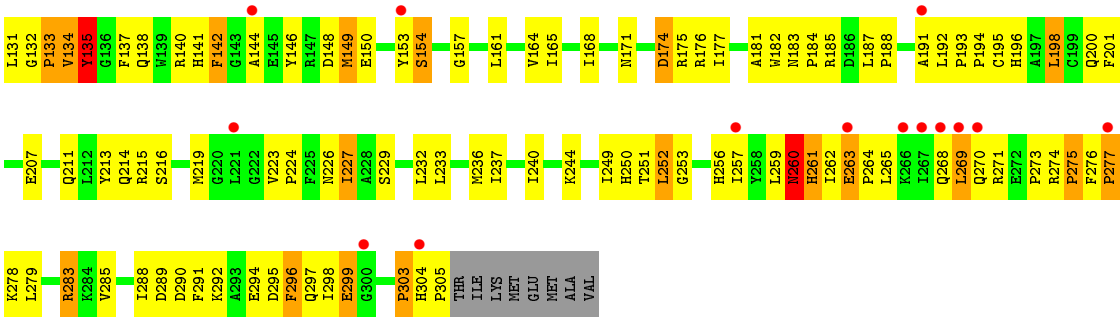


Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total	O	0	0
			16	16		
4	B	12	Total	O	0	0
			12	12		
4	C	12	Total	O	0	0
			12	12		
4	D	8	Total	O	0	0
			8	8		
4	E	16	Total	O	0	0
			16	16		



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	201.80Å 123.11Å 99.93Å 90.00° 115.75° 90.00°	Depositor
Resolution (Å)	45.46 – 2.75 49.86 – 2.75	Depositor EDS
% Data completeness (in resolution range)	78.7 (45.46-2.75) 82.5 (49.86-2.75)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.70 (at 2.73Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.234 , 0.293 0.245 , 0.240	Depositor DCC
R_{free} test set	4772 reflections (10.08%)	DCC
Wilson B-factor (Å ²)	63.4	Xtriage
Anisotropy	0.682	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 65.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 50883 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11445	wwPDB-VP
Average B, all atoms (Å ²)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GSH, PO4

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 $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.47	0/2328	0.70	0/3150
1	B	0.39	0/2328	0.66	0/3150
1	C	0.41	0/2321	0.69	1/3140 (0.0%)
1	D	0.38	0/2328	0.71	3/3150 (0.1%)
1	E	0.37	0/2321	0.67	0/3140
All	All	0.40	0/11626	0.69	4/15730 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
1	D	0	2
All	All	0	3

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	196	HIS	N-CA-C	-8.40	88.32	111.00
1	D	34	LEU	N-CA-C	-5.60	95.88	111.00
1	D	195	CYS	O-C-N	-5.51	113.88	122.70
1	C	198	LEU	CA-CB-CG	5.08	126.98	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	230	TYR	Sidechain
1	D	195	CYS	Mainchain
1	D	213	TYR	Sidechain

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2268	0	2235	99	0
1	B	2268	0	2235	171	0
1	C	2261	0	2228	158	0
1	D	2268	0	2234	235	0
1	E	2261	0	2228	233	0
2	A	20	0	15	0	0
2	D	20	0	12	10	0
3	A	5	0	0	0	0
3	C	5	0	0	0	0
3	E	5	0	0	0	0
4	A	16	0	0	1	0
4	B	12	0	0	1	0
4	C	12	0	0	0	0
4	D	8	0	0	3	0
4	E	16	0	0	1	0
All	All	11445	0	11187	862	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 38.

The worst 5 of 862 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:281:ILE:H	1:D:281:ILE:HD13	1.17	1.04
1:D:177:ILE:HG21	1:D:201:PHE:HB2	1.37	1.03
1:E:263:GLU:HG2	1:E:264:PRO:HD3	1.32	1.03
1:E:118:LEU:H	1:E:118:LEU:HD12	1.24	1.00
1:E:50:ARG:HH11	1:E:50:ARG:HB3	1.27	0.95

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	279/313 (89%)	243 (87%)	30 (11%)	6 (2%)	8	24
1	B	279/313 (89%)	236 (85%)	31 (11%)	12 (4%)	3	9
1	C	278/313 (89%)	231 (83%)	33 (12%)	14 (5%)	3	6
1	D	279/313 (89%)	207 (74%)	52 (19%)	20 (7%)	1	2
1	E	278/313 (89%)	200 (72%)	49 (18%)	29 (10%)	1	1
All	All	1393/1565 (89%)	1117 (80%)	195 (14%)	81 (6%)	2	5

5 of 81 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	54	GLY
1	B	114	SER
1	B	135	TYR
1	C	103	SER
1	D	126	ARG

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	245/271 (90%)	230 (94%)	15 (6%)	23	52
1	B	245/271 (90%)	223 (91%)	22 (9%)	12	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	244/271 (90%)	222 (91%)	22 (9%)	12	31
1	D	245/271 (90%)	212 (86%)	33 (14%)	5	12
1	E	244/271 (90%)	221 (91%)	23 (9%)	11	28
All	All	1223/1355 (90%)	1108 (91%)	115 (9%)	11	28

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	225	PHE
1	D	69	ASP
1	E	252	LEU
1	C	229	SER
1	C	302	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	138	GLN
1	C	268	GLN
1	E	200	GLN
1	C	196	HIS
1	C	205	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

5 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$
3	PO4	A	365	-	4,4,4	1.09	0	6,6,6	0.27	0
2	GSH	A	581	-	13,19,19	0.92	0	15,24,24	1.91	4 (26%)
3	PO4	C	365	-	4,4,4	1.16	0	6,6,6	0.27	0
2	GSH	D	314	1	13,19,19	1.40	2 (15%)	15,24,24	5.63	12 (80%)
3	PO4	E	366	-	4,4,4	1.11	0	6,6,6	0.27	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
3	PO4	A	365	-	-	0/0/0/0	0/0/0/0
2	GSH	A	581	-	-	0/18/24/24	0/0/0/0
3	PO4	C	365	-	-	0/0/0/0	0/0/0/0
2	GSH	D	314	1	-	0/18/24/24	0/0/0/0
3	PO4	E	366	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	314	GSH	CA2-C2	-2.80	1.45	1.52
2	D	314	GSH	CD1-N2	-2.26	1.29	1.34

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	314	GSH	CG1-CD1-N2	-8.94	101.25	115.83
2	D	314	GSH	CB2-CA2-N2	-5.18	104.13	111.40
2	D	314	GSH	CA2-C2-N3	-4.70	107.49	116.72

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	314	GSH	CA2-CB2-SG2	-4.12	109.09	114.16
2	A	581	GSH	C2-CA2-N2	-3.26	102.07	111.26

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	314	GSH	10	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	281/313 (89%)	-0.01	1 (0%) 93 92	40, 65, 98, 114	0
1	B	281/313 (89%)	0.20	13 (4%) 36 29	42, 83, 126, 142	0
1	C	280/313 (89%)	0.24	7 (2%) 61 54	42, 83, 128, 153	0
1	D	281/313 (89%)	0.54	26 (9%) 11 7	49, 110, 137, 159	0
1	E	280/313 (89%)	0.46	17 (6%) 25 18	54, 110, 154, 165	0
All	All	1403/1565 (89%)	0.28	64 (4%) 36 29	40, 88, 136, 165	0

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	144	ALA	4.7
1	C	74	LEU	4.4
1	D	89	LEU	4.1
1	B	123	PHE	3.9
1	B	306	THR	3.8

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	GSH	D	314	20/20	0.69	0.55	9.19	63,102,125,128	20
2	GSH	A	581	20/20	0.87	0.36	5.45	46,118,153,154	0
3	PO4	A	365	5/5	0.96	0.19	0.13	87,89,90,94	0
3	PO4	C	365	5/5	0.91	0.19	-3.26	118,119,120,122	0
3	PO4	E	366	5/5	0.86	0.22	-	148,150,151,152	0

6.5 Other polymers [i](#)

There are no such residues in this entry.