Full wwPDB X-ray Structure Validation Report (i)

Dec 3, 2023 – 10:52 am GMT

PDB ID	:	1GPU
Title	:	Transketolase complex with reaction intermediate
Authors	:	Fiedler, E.; Thorell, S.; Sandalova, T.; Koenig, S.; Schneider, G.
Deposited on	:	2001-11-09
Resolution	:	1.86 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:

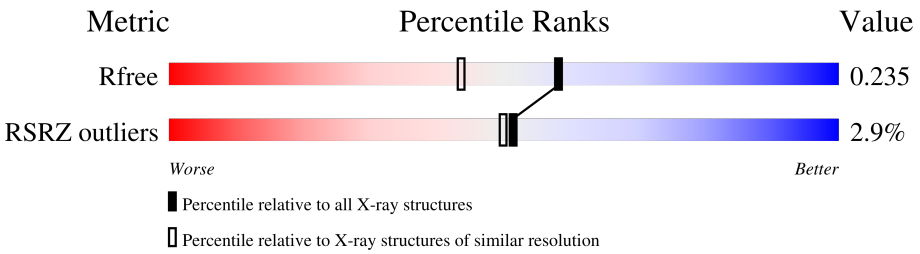
MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.86 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}	130704	2469 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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

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	CA	B	1681	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11315 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.

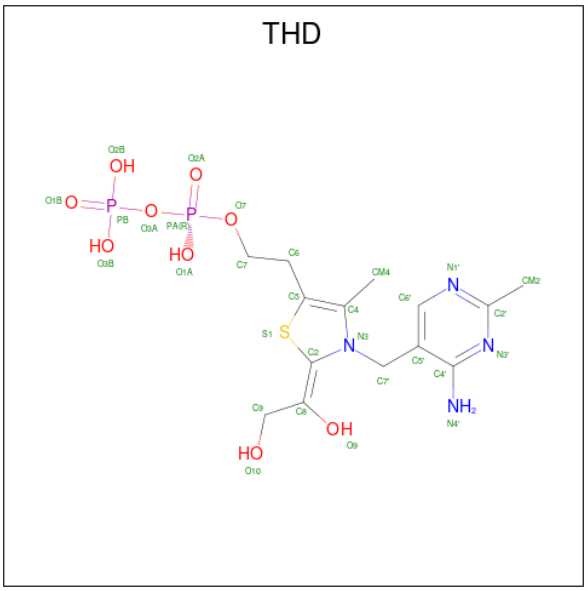
- Molecule 1 is a protein called TRANSKETOLASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	678	5197	3312	884	989	12	0	0	0
1	B	678	5197	3312	884	989	12	0	0	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is 2-[3-[(4-AMINO-2-METHYL-5-PYRIMIDINYL)METHYL]-2-(1,2-DIHYDROXYETHYL)-4-METHYL-1,3-THIAZOL-3-IUM-5-YL]ETHYL TRIHYDROGEN DIPHOSPHATE (three-letter code: THD) (formula: $C_{14}H_{22}N_4O_9P_2S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
3	A	1	30	14	4	9	2	1	0	0
3	B	1	30	14	4	9	2	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	405	405	405	0	0
4	B	454	454	454	0	0

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

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	74.30Å 113.20Å 159.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.86 20.04 – 1.86	Depositor EDS
% Data completeness (in resolution range)	92.8 (30.00-1.86) 92.3 (20.04-1.86)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.06 (at 1.86Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.198 , 0.225 0.208 , 0.235	Depositor DCC
R_{free} test set	5204 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	18.0	Xtriage
Anisotropy	0.841	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 61.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11315	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
3	THD	B	1682	2	28,31,31	2.08	9 (32%)	32,46,46	1.89	12 (37%)
3	THD	A	1682	2	28,31,31	1.98	8 (28%)	32,46,46	1.93	12 (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
3	THD	B	1682	2	-	4/17/23/23	0/2/2/2
3	THD	A	1682	2	-	3/17/23/23	0/2/2/2

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1682	THD	C7'-N3	6.48	1.57	1.47
3	A	1682	THD	C7'-N3	5.97	1.56	1.47
3	B	1682	THD	C4'-N3'	4.25	1.41	1.35
3	A	1682	THD	C4'-N3'	3.94	1.40	1.35
3	B	1682	THD	O9-C8	-2.78	1.25	1.32
3	B	1682	THD	C5-S1	-2.76	1.68	1.74
3	A	1682	THD	O9-C8	-2.73	1.25	1.32
3	B	1682	THD	C6'-N1'	2.69	1.40	1.34
3	A	1682	THD	C6'-N1'	2.53	1.39	1.34
3	B	1682	THD	PB-O2B	-2.32	1.45	1.54
3	A	1682	THD	C5-S1	-2.28	1.69	1.74
3	A	1682	THD	PB-O2B	-2.27	1.46	1.54
3	B	1682	THD	C7'-C5'	2.19	1.55	1.51
3	B	1682	THD	C4-N3	-2.18	1.35	1.39
3	A	1682	THD	C4-N3	-2.18	1.35	1.39
3	A	1682	THD	C5'-C4'	2.17	1.46	1.42
3	B	1682	THD	C5'-C4'	2.14	1.46	1.42

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1682	THD	PA-O3A-PB	3.84	145.99	132.83

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1682	THD	PA-O3A-PB	3.74	145.67	132.83
3	A	1682	THD	C7'-N3-C4	-3.54	120.56	125.01
3	B	1682	THD	C5'-C7'-N3	-3.51	108.67	113.13
3	B	1682	THD	C7'-N3-C4	-3.25	120.92	125.01
3	A	1682	THD	O7-PA-O2A	-3.24	96.42	109.07
3	B	1682	THD	O7-PA-O2A	-3.17	96.70	109.07
3	A	1682	THD	C5'-C7'-N3	-2.94	109.39	113.13
3	B	1682	THD	N1'-C2'-N3'	-2.83	120.66	125.54
3	A	1682	THD	C7'-C5'-C4'	2.83	125.43	122.56
3	A	1682	THD	O3A-PB-O1B	-2.82	95.56	111.19
3	B	1682	THD	O3A-PB-O1B	-2.77	95.85	111.19
3	A	1682	THD	N1'-C2'-N3'	-2.75	120.81	125.54
3	A	1682	THD	C6'-N1'-C2'	2.60	120.39	115.96
3	B	1682	THD	C6'-N1'-C2'	2.60	120.39	115.96
3	A	1682	THD	CM2-C2'-N3'	2.33	120.78	117.15
3	B	1682	THD	CM2-C2'-N3'	2.30	120.74	117.15
3	B	1682	THD	C7'-C5'-C4'	2.28	124.87	122.56
3	A	1682	THD	O3B-PB-O1B	2.21	119.33	110.68
3	B	1682	THD	O3B-PB-O1B	2.19	119.26	110.68
3	A	1682	THD	O9-C8-C9	2.14	119.46	113.91
3	B	1682	THD	O9-C8-C9	2.10	119.36	113.91
3	A	1682	THD	O1A-PA-O7	2.09	117.45	107.75
3	B	1682	THD	O1A-PA-O7	2.03	117.15	107.75

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1682	THD	PB-O3A-PA-O7
3	B	1682	THD	PB-O3A-PA-O7
3	A	1682	THD	C5-C6-C7-O7
3	B	1682	THD	C5-C6-C7-O7
3	B	1682	THD	S1-C5-C6-C7
3	A	1682	THD	S1-C5-C6-C7
3	B	1682	THD	C4-C5-C6-C7

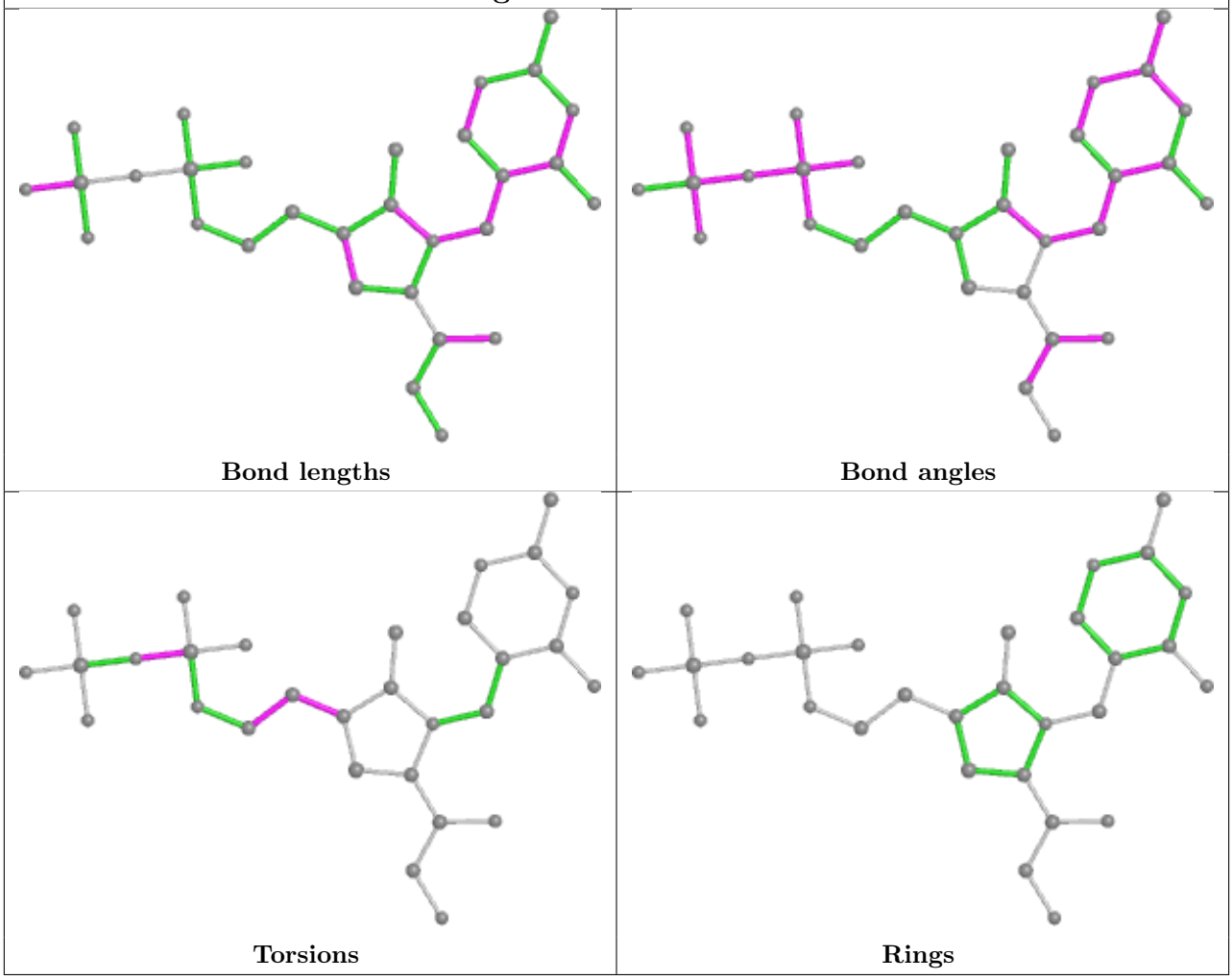
There are no ring outliers.

No monomer is involved in short contacts.

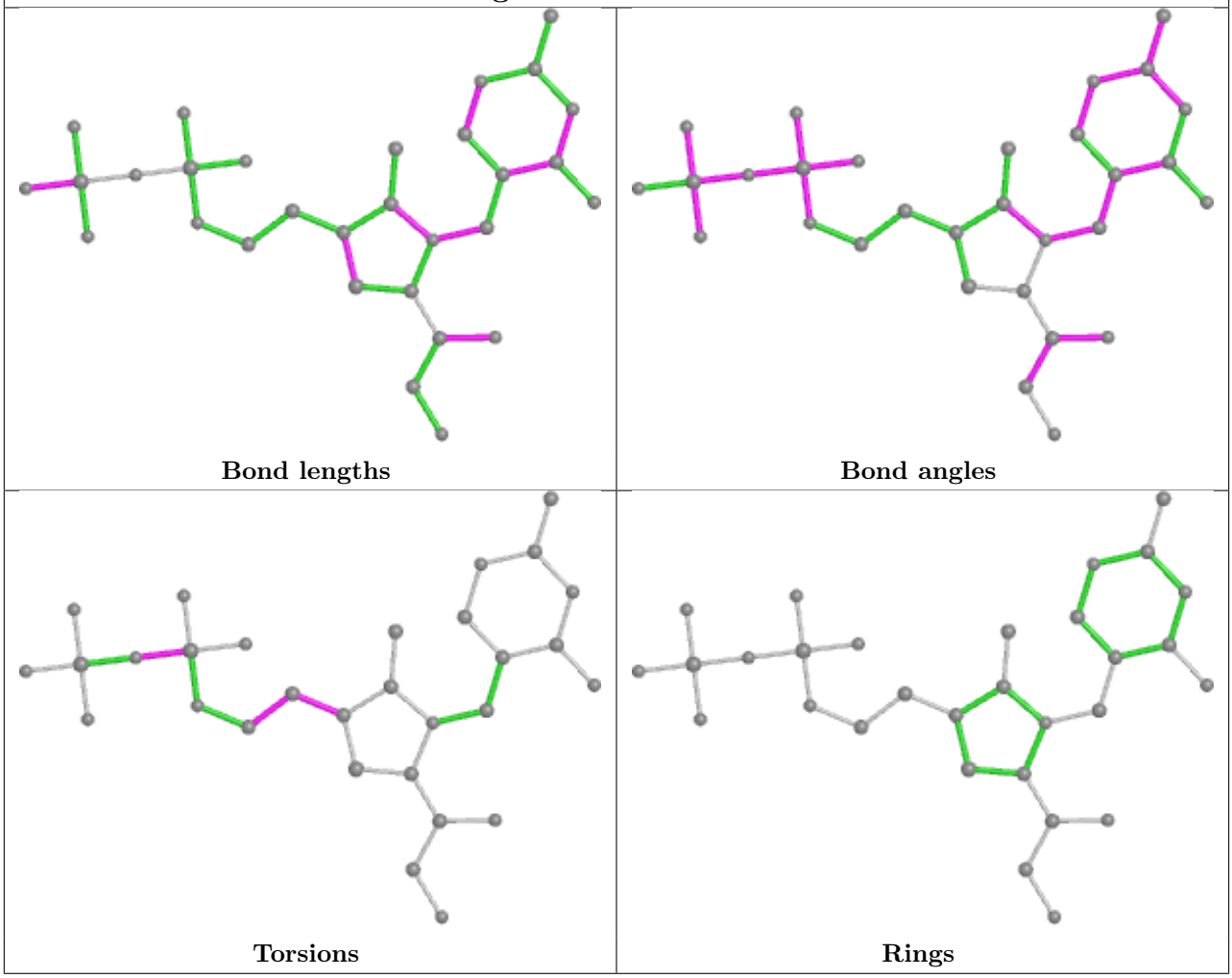
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

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 THD B 1682



Ligand THD A 1682



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)

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	678/680 (99%)	0.19	27 (3%)	38	36	13, 20, 32, 47	0
1	B	678/680 (99%)	0.07	13 (1%)	66	66	12, 18, 31, 46	0
All	All	1356/1360 (99%)	0.13	40 (2%)	51	50	12, 19, 32, 47	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	528	ARG	5.9
1	A	402	SER	4.1
1	A	393	GLU	3.9
1	A	3	GLN	3.6
1	A	341	ASN	3.4
1	A	323	PHE	3.4
1	B	554	ASN	3.1
1	B	528	ARG	3.0
1	A	340	ALA	3.0
1	A	401	SER	2.8
1	B	314	LYS	2.8
1	A	580	ASN	2.8
1	B	53	THR	2.7
1	B	392	LYS	2.7
1	A	236	LEU	2.6
1	A	352	LYS	2.6
1	B	352	LYS	2.6
1	A	669	GLY	2.6
1	B	3	GLN	2.5
1	A	280	GLY	2.4
1	A	371	ASN	2.4
1	A	264	GLY	2.3
1	A	554	ASN	2.3
1	B	598	PRO	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	345	LYS	2.3
1	A	337	GLN	2.2
1	A	322	LYS	2.2
1	A	325	GLU	2.2
1	B	325	GLU	2.2
1	A	283	PRO	2.1
1	B	535	GLU	2.1
1	B	660	ALA	2.1
1	A	391	TRP	2.1
1	A	321	LYS	2.1
1	A	518	LYS	2.1
1	A	404	SER	2.1
1	A	405	GLY	2.1
1	B	143	GLY	2.1
1	B	574	LYS	2.0
1	A	372	GLN	2.0

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

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

5.3 Carbohydrates (i)

There are no monosaccharides in this entry.

5.4 Ligands (i)

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. 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.

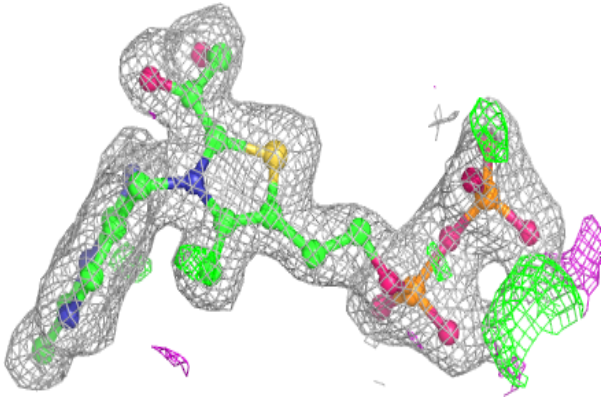
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CA	B	1681	1/1	0.46	0.41	71,71,71,71	0
3	THD	B	1682	30/30	0.95	0.10	11,19,24,24	0
3	THD	A	1682	30/30	0.96	0.09	13,20,23,26	0
2	CA	A	1681	1/1	0.99	0.05	18,18,18,18	0

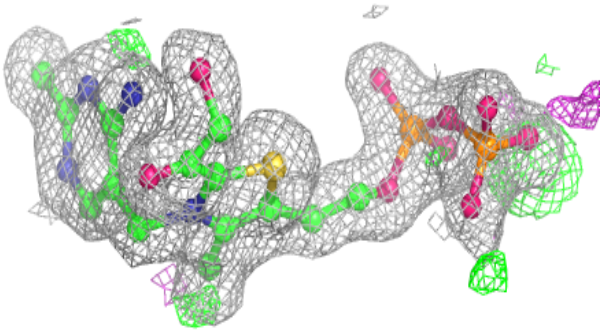
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

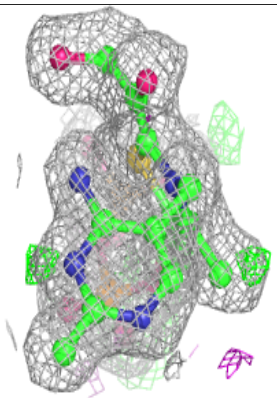
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around THD B 1682:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

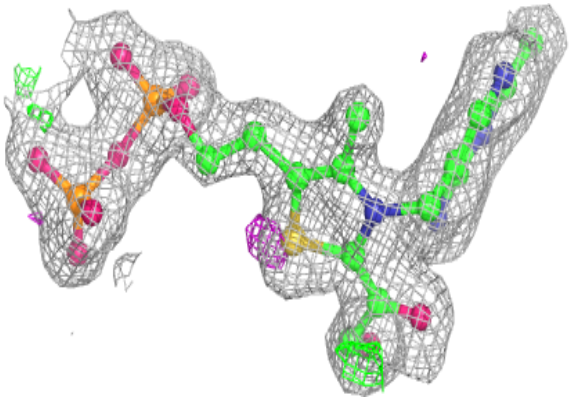


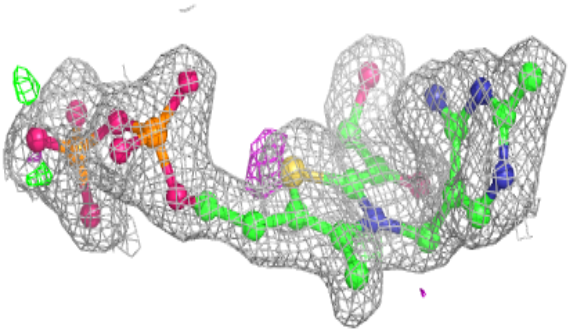


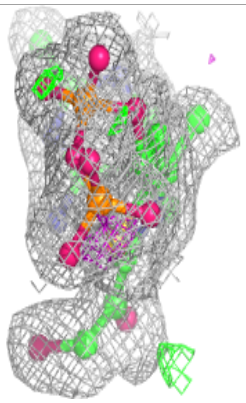


Electron density around THD A 1682:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







5.5 Other polymers (i)

There are no such residues in this entry.