



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 13, 2025 – 12:13 PM EDT

PDB ID : 9E97 / pdb_00009e97
Title : L-allo-threonine aldolase from Thermotoga maritima N308E-Y87A-R122G-P
121D Mutant with a 2-(aminomethyl)pyridine PLP modification
Authors : Wang, S.; Jeffrey, P.D.; Sorigue, D.; Hyster, T.K.
Deposited on : 2024-11-07
Resolution : 1.53 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.44

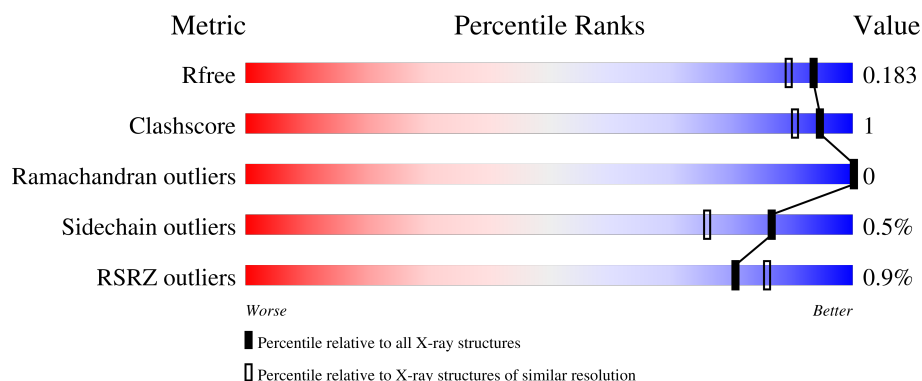
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

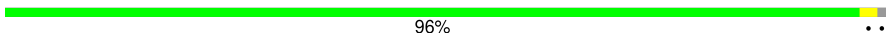

The reported resolution of this entry is 1.53 Å.

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}	164625	3511 (1.56-1.52)
Clashscore	180529	3784 (1.56-1.52)
Ramachandran outliers	177936	3720 (1.56-1.52)
Sidechain outliers	177891	3717 (1.56-1.52)
RSRZ outliers	164620	3510 (1.56-1.52)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	349	 96%
1	B	349	 94%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5830 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 L-allo-threonine aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	344	Total	C	N	O	S	0	13	0
			2700	1692	485	505	18			
1	B	344	Total	C	N	O	S	0	9	0
			2676	1674	478	506	18			

There are 28 discrepancies between the modelled and reference sequences:

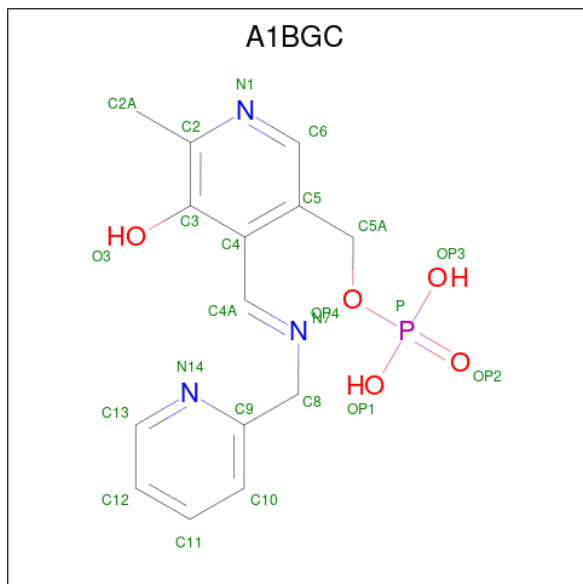
Chain	Residue	Modelled	Actual	Comment	Reference
A	87	ALA	TYR	engineered mutation	UNP Q9X266
A	121	ASP	PRO	engineered mutation	UNP Q9X266
A	122	GLY	ARG	engineered mutation	UNP Q9X266
A	308	GLU	ASN	engineered mutation	UNP Q9X266
A	340	SER	ARG	engineered mutation	UNP Q9X266
A	341	GLN	LYS	engineered mutation	UNP Q9X266
A	342	VAL	-	expression tag	UNP Q9X266
A	343	GLN	-	expression tag	UNP Q9X266
A	344	HIS	-	expression tag	UNP Q9X266
A	345	HIS	-	expression tag	UNP Q9X266
A	346	HIS	-	expression tag	UNP Q9X266
A	347	HIS	-	expression tag	UNP Q9X266
A	348	HIS	-	expression tag	UNP Q9X266
A	349	HIS	-	expression tag	UNP Q9X266
B	87	ALA	TYR	engineered mutation	UNP Q9X266
B	121	ASP	PRO	engineered mutation	UNP Q9X266
B	122	GLY	ARG	engineered mutation	UNP Q9X266
B	308	GLU	ASN	engineered mutation	UNP Q9X266
B	340	SER	ARG	engineered mutation	UNP Q9X266
B	341	GLN	LYS	engineered mutation	UNP Q9X266
B	342	VAL	-	expression tag	UNP Q9X266
B	343	GLN	-	expression tag	UNP Q9X266
B	344	HIS	-	expression tag	UNP Q9X266
B	345	HIS	-	expression tag	UNP Q9X266
B	346	HIS	-	expression tag	UNP Q9X266

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	347	HIS	-	expression tag	UNP Q9X266
B	348	HIS	-	expression tag	UNP Q9X266
B	349	HIS	-	expression tag	UNP Q9X266

- Molecule 2 is {5-hydroxy-6-methyl-4-[(E)-{[(pyridin-2-yl)methyl]imino}methyl]pyridin-3-yl}methyl dihydrogen phosphate (CCD ID: A1BGC) (formula: C₁₄H₁₆N₃O₅P) (labeled as "Ligand of Interest" by depositor).



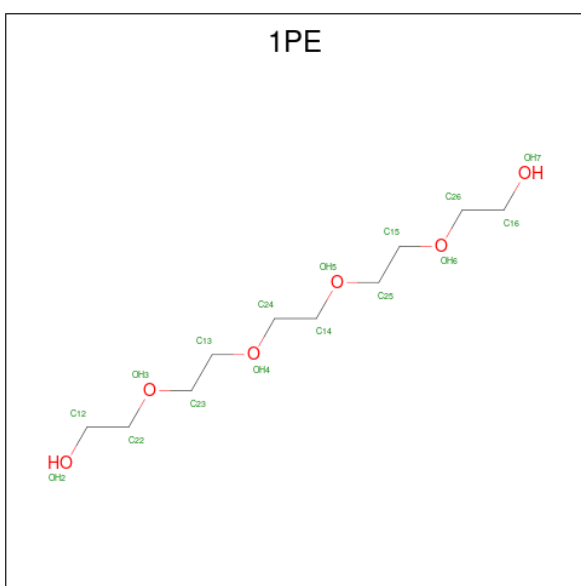
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			23	14	3	5	1		
2	B	1	Total	C	N	O	P	0	0
			23	14	3	5	1		

- Molecule 3 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	8	5		
3	B	1	Total	C	O	0	0
			13	8	5		

- Molecule 4 is PENTAETHYLENE GLYCOL (CCD ID: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	10	6		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 6 is CALCIUM ION (CCD ID: CA) (formula: Ca).

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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	189	Total	O	0	0
			189	189		
7	B	152	Total	O	0	0
			152	152		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

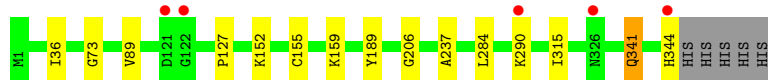
- Molecule 1: L-allo-threonine aldolase

Chain A:  96% ..



- Molecule 1: L-allo-threonine aldolase

Chain B:  94% ..



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	94.21Å 166.02Å 93.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.56 – 1.53 29.56 – 1.53	Depositor EDS
% Data completeness (in resolution range)	96.6 (29.56-1.53) 95.7 (29.56-1.53)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 1.54Å)	Xtriage
Refinement program	PHENIX 1.17_3644	Depositor
R, R_{free}	0.169 , 0.186 0.167 , 0.183	Depositor DCC
R_{free} test set	5401 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	14.8	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 38.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.001 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5830	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.58 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.6768e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 1PE, CA, A1BGC, PG4, PEG

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	0.29	0/2772	0.52	0/3740
1	B	0.27	0/2736	0.53	1/3692 (0.0%)
All	All	0.28	0/5508	0.52	1/7432 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	206	GLY	N-CA-C	8.65	120.17	111.95

There are no chirality outliers.

There are no planarity outliers.

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	2700	0	2765	6	0
1	B	2676	0	2719	7	0
2	A	23	0	0	0	0
2	B	23	0	0	0	0
3	A	13	0	17	0	0
3	B	13	0	17	0	0
4	A	16	0	22	3	0
5	A	14	0	20	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	7	0	10	0	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	A	189	0	0	0	0
7	B	152	0	0	0	0
All	All	5830	0	5570	13	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 1.

The worst 5 of 13 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:A:155:CYS:SG	1:A:159:LYS:NZ	2.55	0.71
1:B:290:LYS:HE3	1:B:344:HIS:HA	1.79	0.64
1:A:290:LYS:HE2	4:A:403:1PE:H242	1.82	0.61
1:B:36:ILE:HD13	1:B:237:ALA:HB2	1.94	0.50
1:B:155:CYS:SG	1:B:159:LYS:HE3	2.53	0.48

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	355/349 (102%)	347 (98%)	8 (2%)	0	100	100
1	B	351/349 (101%)	343 (98%)	8 (2%)	0	100	100
All	All	706/698 (101%)	690 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	292/285 (102%)	291 (100%)	1 (0%)	91	84
1	B	288/285 (101%)	286 (99%)	2 (1%)	81	66
All	All	580/570 (102%)	577 (100%)	3 (0%)	86	76

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	VAL
1	B	89	VAL
1	B	341	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	164	ASN
1	A	174	ASN
1	B	174	ASN
1	B	63	GLN
1	A	150	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 4 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	PEG	A	405	-	6,6,6	0.13	0	5,5,5	0.14	0
5	PEG	B	403	-	6,6,6	0.18	0	5,5,5	0.15	0
2	A1BGC	A	401	-	24,24,24	0.38	0	31,33,33	0.63	0
3	PG4	A	402	6	12,12,12	0.13	0	11,11,11	0.25	0
3	PG4	B	402	6	12,12,12	0.13	0	11,11,11	0.23	0
2	A1BGC	B	401	-	24,24,24	0.39	0	31,33,33	0.73	1 (3%)
5	PEG	A	404	-	6,6,6	0.15	0	5,5,5	0.17	0
4	1PE	A	403	-	15,15,15	0.16	0	14,14,14	0.18	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
5	PEG	A	405	-	-	0/4/4/4	-
5	PEG	B	403	-	-	2/4/4/4	-
2	A1BGC	A	401	-	-	3/12/12/12	0/2/2/2
3	PG4	A	402	6	-	0/10/10/10	-
3	PG4	B	402	6	-	1/10/10/10	-
2	A1BGC	B	401	-	-	3/12/12/12	0/2/2/2
5	PEG	A	404	-	-	2/4/4/4	-
4	1PE	A	403	-	-	3/13/13/13	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	A1BGC	OP4-C5A-C5	2.04	113.18	109.36

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

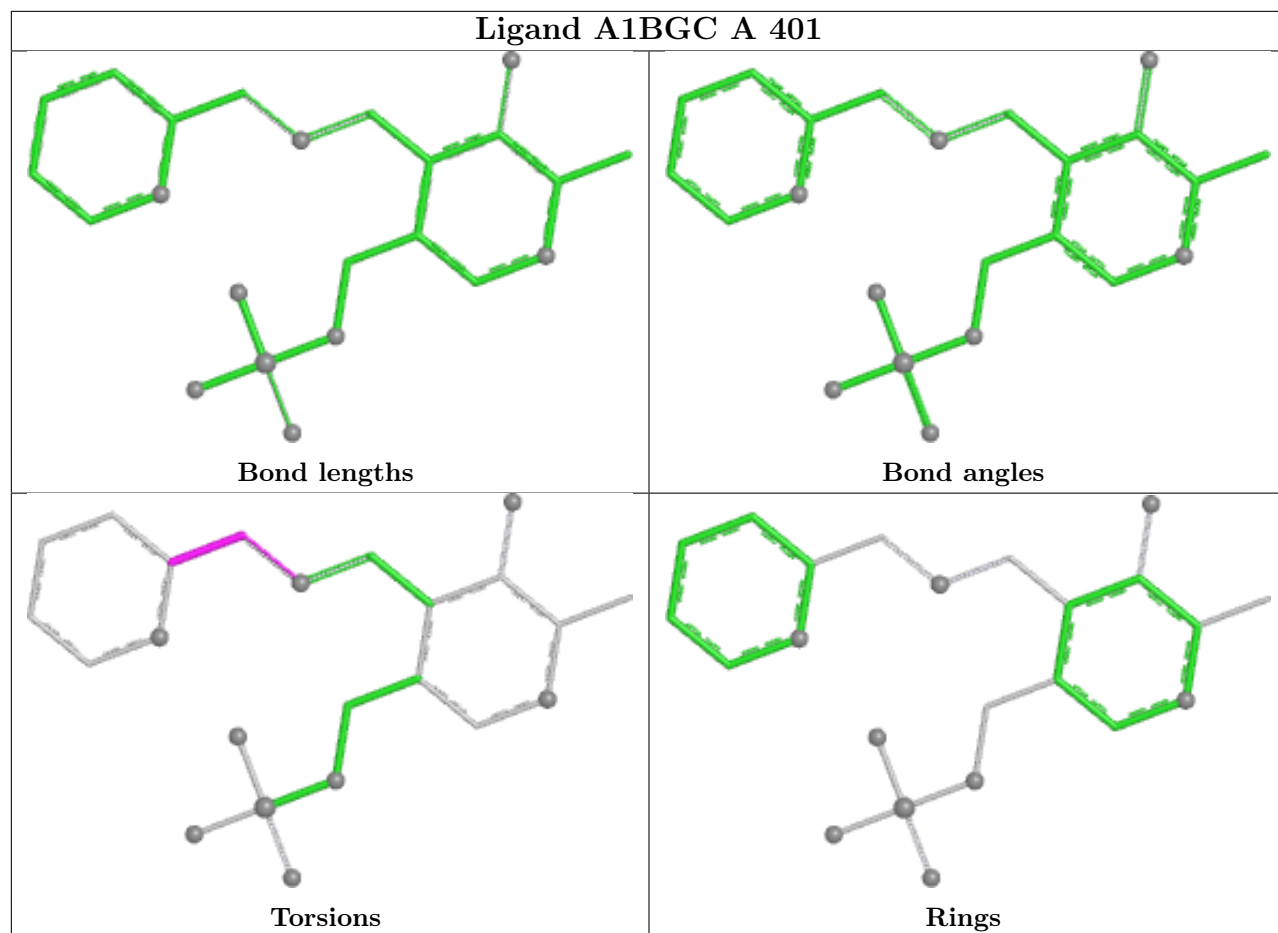
Mol	Chain	Res	Type	Atoms
5	A	404	PEG	O2-C3-C4-O4
4	A	403	1PE	C13-C23-OH3-C22
2	A	401	A1BGC	C9-C8-N7-C4A
2	B	401	A1BGC	C9-C8-N7-C4A
5	B	403	PEG	C4-C3-O2-C2

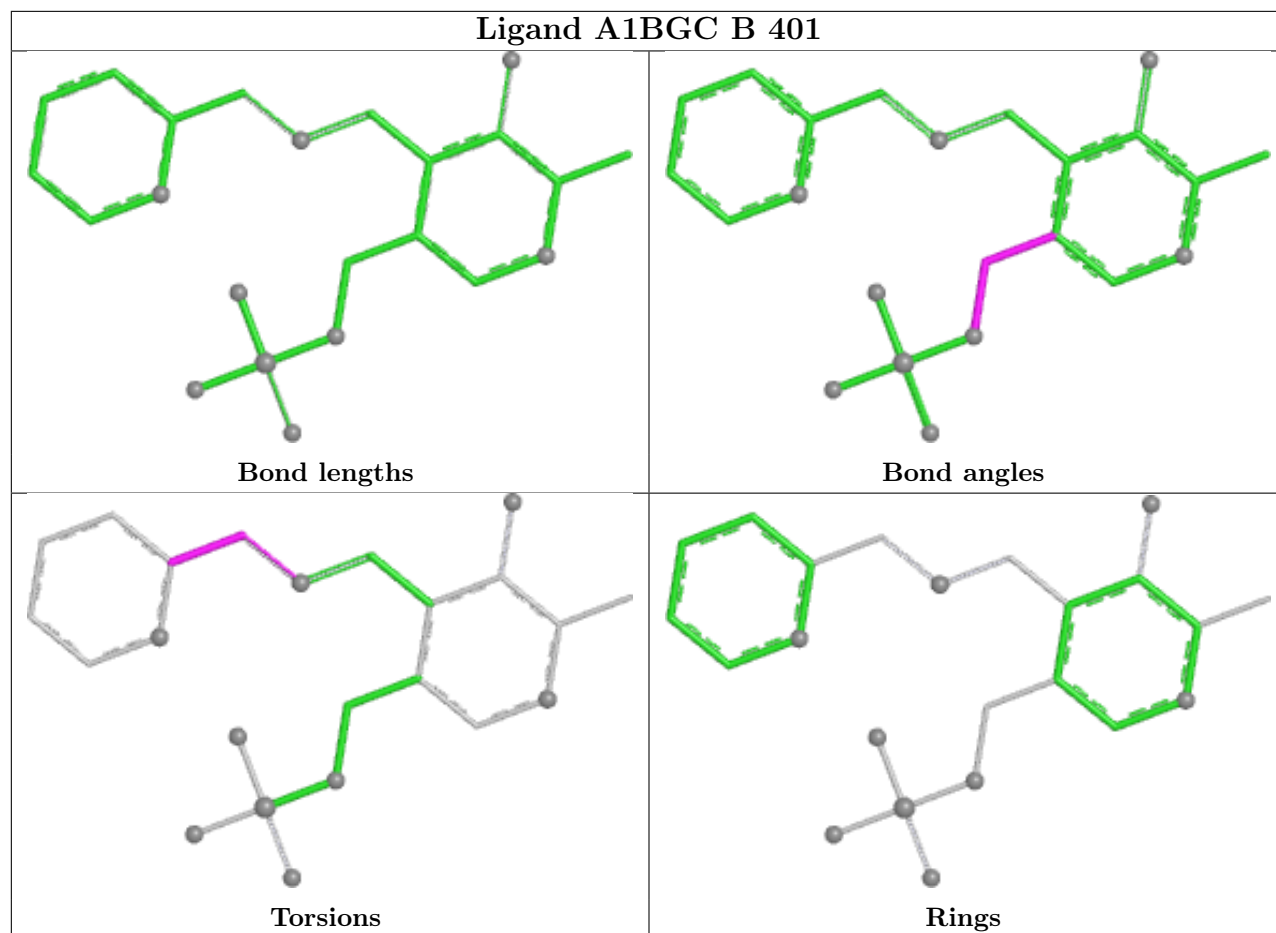
There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	403	1PE	3	0

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 than 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.





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 [i](#)

6.1 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	344/349 (98%)	-0.24	1 (0%)	90 93	7, 16, 30, 49	13 (3%)
1	B	344/349 (98%)	-0.00	5 (1%)	71 79	8, 18, 38, 69	9 (2%)
All	All	688/698 (98%)	-0.12	6 (0%)	81 86	7, 17, 35, 69	22 (3%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	344	HIS	3.5
1	A	344	HIS	3.4
1	B	121	ASP	2.8
1	B	326	ASN	2.1
1	B	122[A]	GLY	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

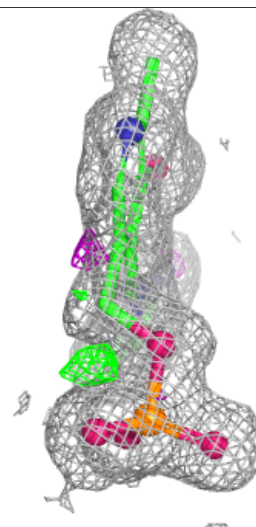
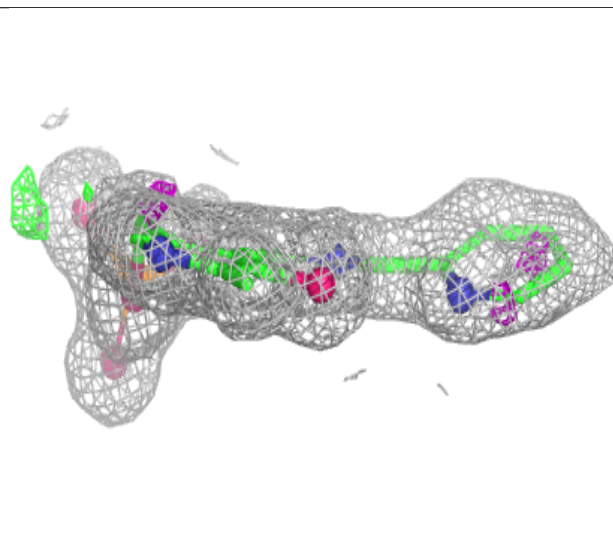
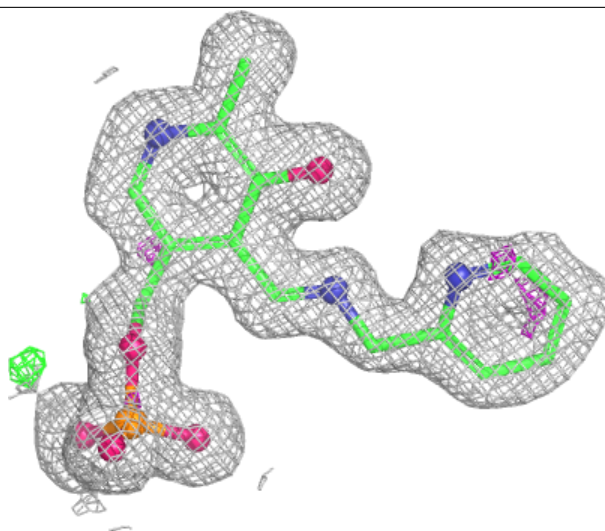
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	PEG	A	404	7/7	0.73	0.18	33,36,38,40	0
4	1PE	A	403	16/16	0.74	0.18	36,41,53,62	0
5	PEG	B	403	7/7	0.74	0.17	23,28,43,44	0
5	PEG	A	405	7/7	0.80	0.17	33,35,44,45	0
3	PG4	B	402	13/13	0.85	0.12	25,34,44,54	0
3	PG4	A	402	13/13	0.89	0.11	22,27,36,53	0
6	CA	B	405	1/1	0.96	0.15	35,35,35,35	0
2	A1BGC	B	401	23/23	0.97	0.07	12,15,23,24	0
6	CA	A	407	1/1	0.98	0.11	27,27,27,27	0
2	A1BGC	A	401	23/23	0.98	0.06	11,13,24,27	0
6	CA	B	404	1/1	0.99	0.06	16,16,16,16	0
6	CA	A	406	1/1	1.00	0.05	14,14,14,14	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.

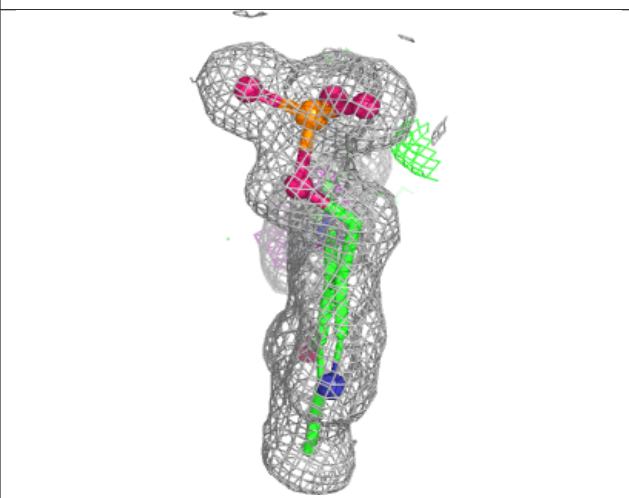
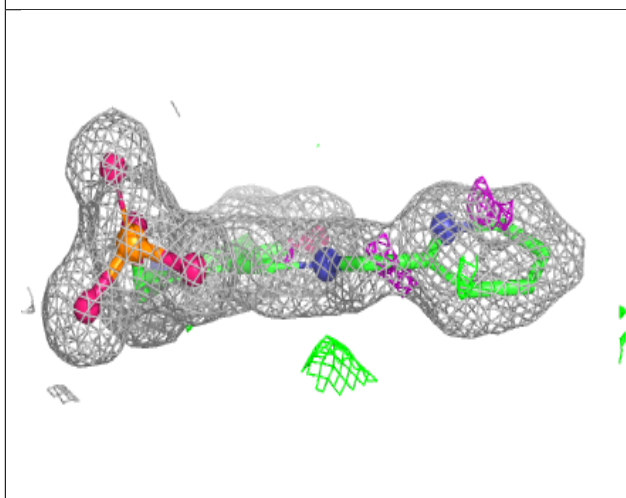
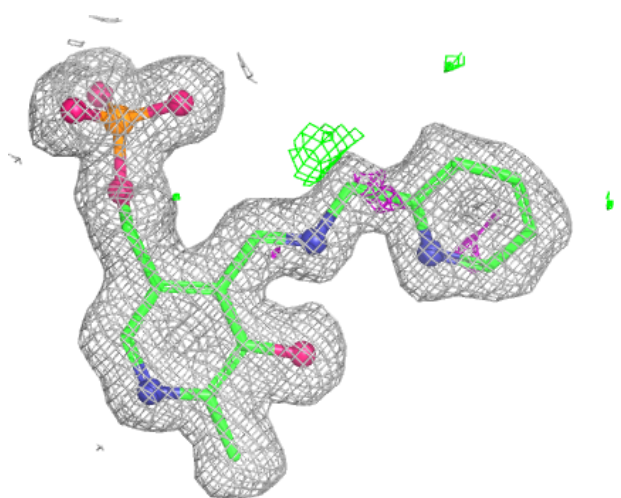
Electron density around A1BGC B 401:

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



Electron density around A1BGC A 401:

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



6.5 Other polymers [i](#)

There are no such residues in this entry.