wwPDB X-ray Structure Validation Summary Report (i)

Apr 21, 2024 – 01:32 pm BST

PDB ID	:	7AIA
Title	:	Complex of human GDAP1 with hexadecanedioic acid
Authors	:	Nguyen, G.T.T.; Sutinen, A.; Kursula, P.
Deposited on	:	2020-09-26
Resolution	:	2.20 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 (i)) were used in the production of this report:

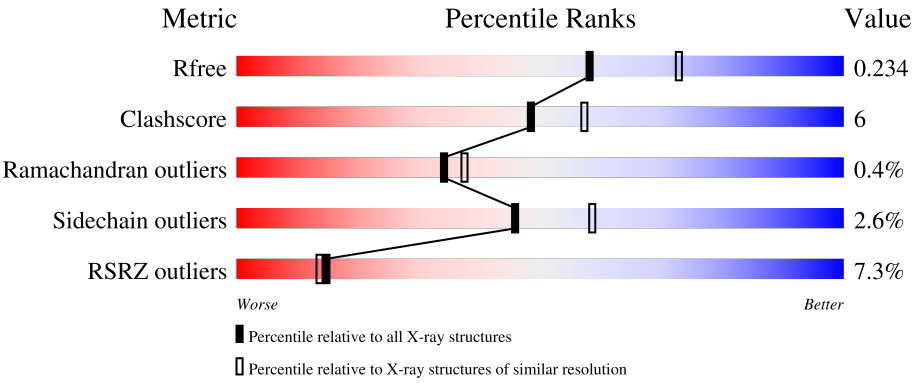
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	AAA	280	6% 79% 11% • 8%
1	BBB	280	7% 75% 12% • 12%

2 Entry composition (i)

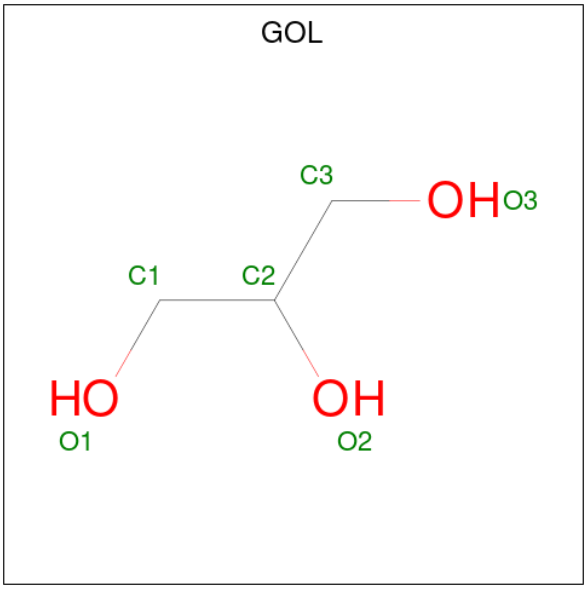
There are 5 unique types of molecules in this entry. The entry contains 8806 atoms, of which 4326 are hydrogens and 0 are deuteriums.

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 Ganglioside-induced differentiation-associated protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	259	Total	C	H	N	O	S	67	2	0
			4333	1377	2176	377	394	9			
1	BBB	245	Total	C	H	N	O	S	63	1	0
			4075	1298	2038	355	375	9			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



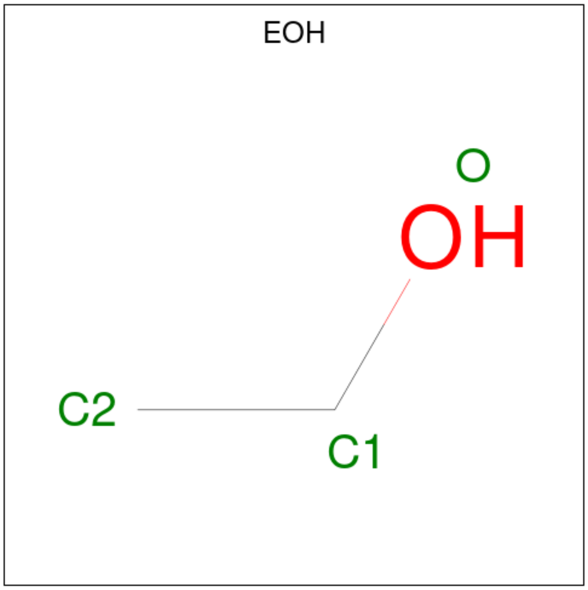
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

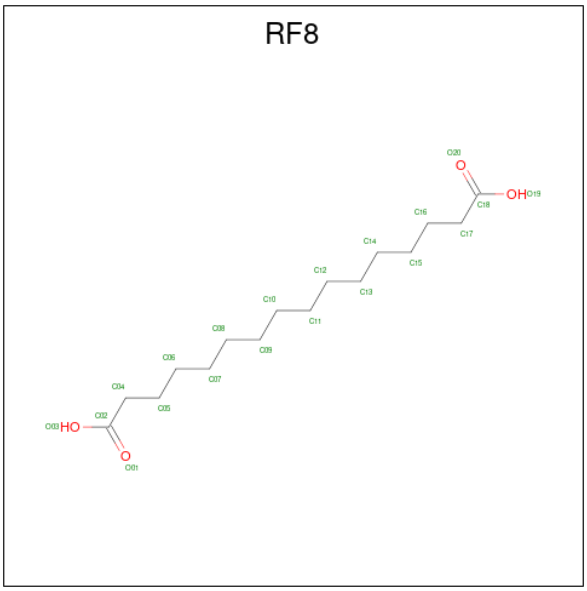
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C H O 14 3 8 3	2	0
2	AAA	1	Total C H O 14 3 8 3	2	0

- Molecule 3 is ETHANOL (three-letter code: EOH) (formula: C_2H_6O).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C H O 9 2 6 1	0	0
3	AAA	1	Total C H O 9 2 6 1	0	0
3	AAA	1	Total C H O 9 2 6 1	0	0
3	AAA	1	Total C H O 9 2 6 1	0	0
3	AAA	1	Total C H O 9 2 6 1	0	0
3	AAA	1	Total C H O 9 2 6 1	0	0

- Molecule 4 is hexadecanedioic acid (three-letter code: RF8) (formula: $C_{16}H_{30}O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total C H O 48 16 28 4	0	0

- Molecule 5 is water.

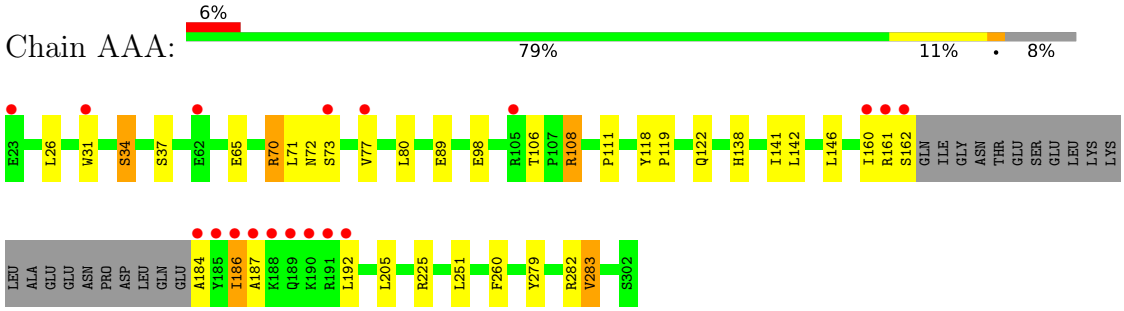
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	134	Total O 134 134	0	0
5	BBB	78	Total O 78 78	0	0

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.

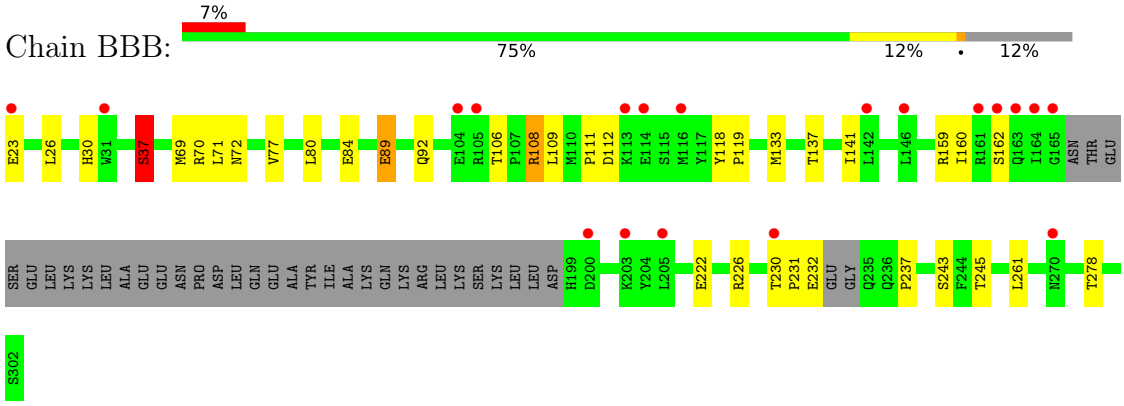
• Molecule 1: Ganglioside-induced differentiation-associated protein 1

Chain AAA:



• Molecule 1: Ganglioside-induced differentiation-associated protein 1

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.05Å 114.91Å 116.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.63 – 2.20 45.63 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.63-2.20) 99.8 (45.63-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.54 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.202 , 0.230 0.209 , 0.234	Depositor DCC
R_{free} test set	2011 reflections (3.98%)	wwPDB-VP
Wilson B-factor (Å^2)	53.9	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 51.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8806	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, EOH, RF8

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	AAA	0.83	2/2213 (0.1%)	0.97	6/2994 (0.2%)
1	BBB	0.73	1/2084 (0.0%)	0.88	3/2822 (0.1%)
All	All	0.79	3/4297 (0.1%)	0.93	9/5816 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	98	GLU	CD-OE2	-5.50	1.19	1.25
1	AAA	65	GLU	CD-OE2	-5.39	1.19	1.25
1	BBB	37	SER	CB-OG	5.26	1.49	1.42

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	161	ARG	CG-CD-NE	10.12	133.04	111.80
1	AAA	225	ARG	NE-CZ-NH2	10.00	125.30	120.30
1	AAA	225	ARG	NE-CZ-NH1	-8.41	116.09	120.30
1	AAA	70	ARG	NE-CZ-NH1	7.33	123.97	120.30
1	AAA	108	ARG	NE-CZ-NH2	6.58	123.59	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	AAA	2157	2176	2162	29	0
1	BBB	2037	2038	2021	24	0
2	AAA	36	48	48	1	0
3	AAA	18	36	36	0	0
4	AAA	20	28	0	1	0
5	AAA	134	0	0	10	0
5	BBB	78	0	0	10	0
All	All	4480	4326	4267	51	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 6.

The worst 5 of 51 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:AAA:279:TYR:CE1	5:AAA:721:HOH:O	2.19	0.93
1:AAA:138[A]:HIS:CD2	1:AAA:142:LEU:CD1	2.71	0.72
1:AAA:283:VAL:HB	5:AAA:721:HOH:O	1.93	0.67
1:AAA:31[A]:TRP:HE3	5:AAA:699:HOH:O	1.76	0.67
1:BBB:77:VAL:HG11	1:BBB:89:GLU:HG2	1.77	0.66

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	AAA	257/280 (92%)	252 (98%)	5 (2%)	0	100 100
1	BBB	240/280 (86%)	234 (98%)	4 (2%)	2 (1%)	19 19
All	All	497/560 (89%)	486 (98%)	9 (2%)	2 (0%)	34 37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	231	PRO
1	BBB	89	GLU

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	AAA	241/258 (93%)	234 (97%)	7 (3%)	42 54
1	BBB	228/258 (88%)	223 (98%)	5 (2%)	52 65
All	All	469/516 (91%)	457 (97%)	12 (3%)	46 58

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	23	GLU
1	BBB	37	SER
1	BBB	162	SER
1	BBB	69	MET
1	AAA	162	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

13 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	EOH	AAA	507	-	2,2,2	0.32	0	1,1,1	0.15	0
2	GOL	AAA	505	-	5,5,5	0.19	0	5,5,5	0.44	0
2	GOL	AAA	511	-	5,5,5	0.29	0	5,5,5	0.71	0
3	EOH	AAA	504	-	2,2,2	0.20	0	1,1,1	0.10	0
3	EOH	AAA	509	-	2,2,2	0.37	0	1,1,1	0.31	0
3	EOH	AAA	508	-	2,2,2	0.50	0	1,1,1	0.25	0
3	EOH	AAA	512	-	2,2,2	0.11	0	1,1,1	0.10	0
2	GOL	AAA	506	-	5,5,5	0.18	0	5,5,5	0.48	0
4	RF8	AAA	513	-	19,19,19	0.98	1 (5%)	20,20,20	0.98	1 (5%)
2	GOL	AAA	502	-	5,5,5	0.18	0	5,5,5	0.51	0
3	EOH	AAA	510	-	2,2,2	0.45	0	1,1,1	0.23	0
2	GOL	AAA	501	-	5,5,5	0.09	0	5,5,5	0.30	0
2	GOL	AAA	503	-	5,5,5	0.13	0	5,5,5	0.29	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
2	GOL	AAA	505	-	-	2/4/4/4	-
2	GOL	AAA	511	-	-	4/4/4/4	-
2	GOL	AAA	506	-	-	4/4/4/4	-
4	RF8	AAA	513	-	-	9/17/17/17	-
2	GOL	AAA	502	-	-	2/4/4/4	-
2	GOL	AAA	501	-	-	2/4/4/4	-
2	GOL	AAA	503	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	AAA	513	RF8	O01-C02	2.16	1.29	1.22

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AAA	513	RF8	C06-C05-C04	-2.05	105.83	113.19

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	502	GOL	C1-C2-C3-O3
2	AAA	503	GOL	O1-C1-C2-C3
2	AAA	505	GOL	C1-C2-C3-O3
2	AAA	511	GOL	O1-C1-C2-C3
2	AAA	511	GOL	C1-C2-C3-O3

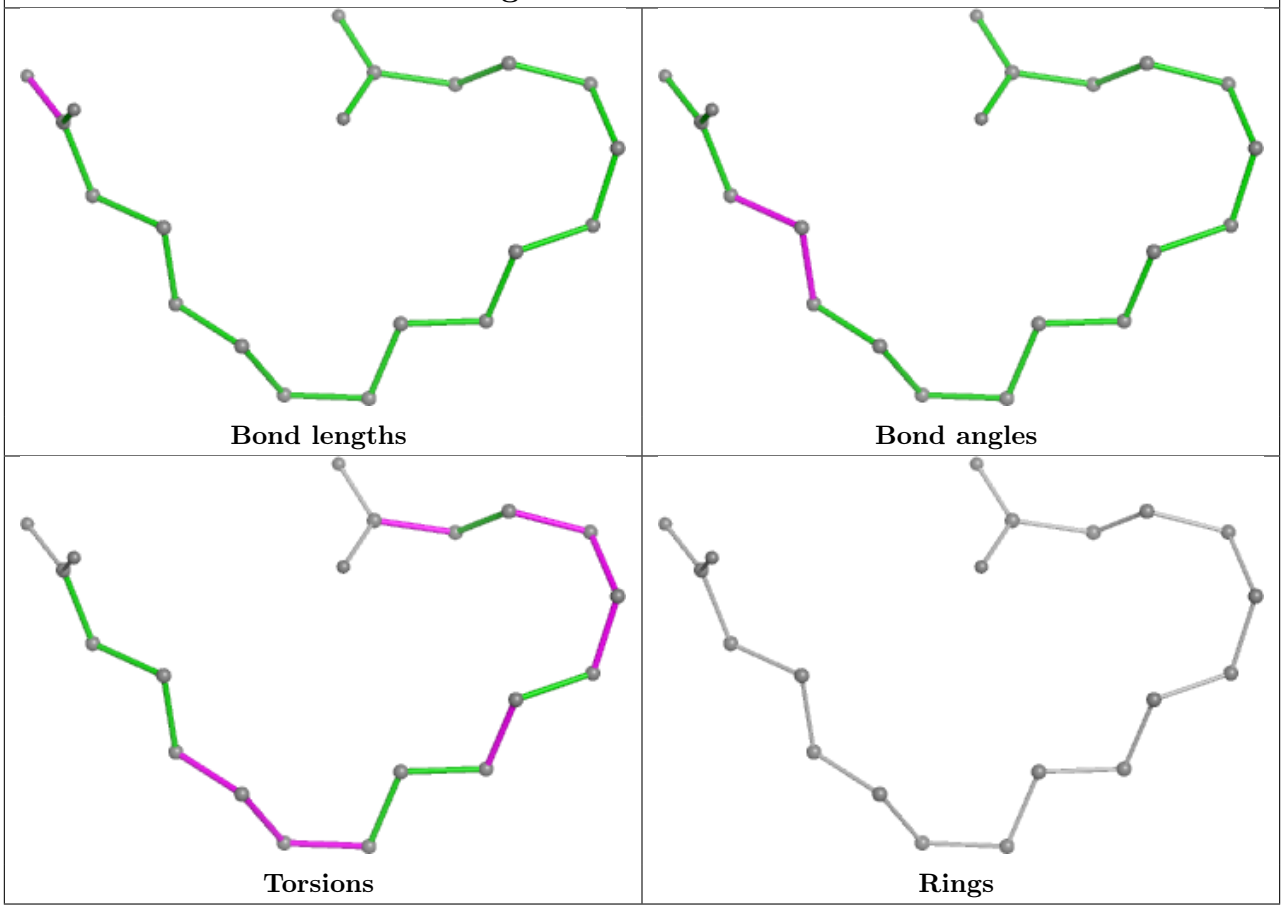
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	511	GOL	1	0
4	AAA	513	RF8	1	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 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 RF8 AAA 513



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, 95^{th} 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	AAA	259/280 (92%)	0.42	18 (6%)	16	15	37, 54, 129, 174	0
1	BBB	245/280 (87%)	0.55	19 (7%)	13	11	48, 80, 143, 182	0
All	All	504/560 (90%)	0.49	37 (7%)	15	14	37, 66, 138, 182	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	165	GLY	7.8
1	AAA	187	ALA	7.0
1	AAA	184	ALA	6.8
1	AAA	185	TYR	6.4
1	BBB	164	ILE	5.3

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 monosaccharides in this entry.

6.4 Ligands

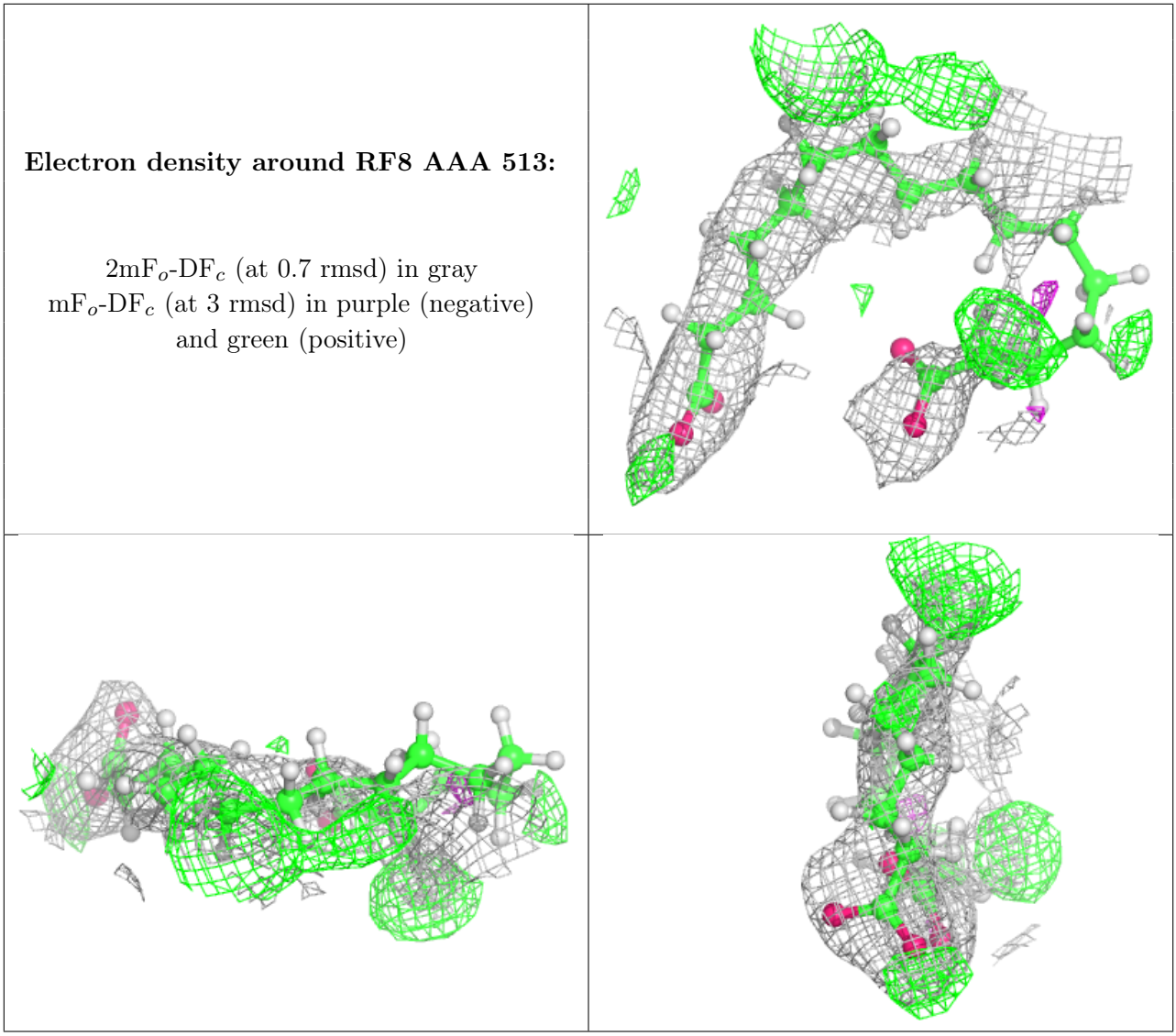
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	EOH	AAA	508	3/3	0.51	0.23	73,77,80,85	0
3	EOH	AAA	504	3/3	0.59	0.30	83,86,90,95	0
3	EOH	AAA	509	3/3	0.68	0.33	70,75,80,89	0
2	GOL	AAA	511	6/6	0.72	0.24	48,86,98,106	2
4	RF8	AAA	513	20/20	0.72	0.31	72,119,143,146	0
2	GOL	AAA	501	6/6	0.74	0.17	48,81,83,85	2
3	EOH	AAA	510	3/3	0.78	0.12	64,78,80,86	0
2	GOL	AAA	505	6/6	0.79	0.29	48,86,94,98	2
3	EOH	AAA	507	3/3	0.79	0.36	61,73,74,83	0
2	GOL	AAA	506	6/6	0.86	0.12	48,96,101,103	2
2	GOL	AAA	503	6/6	0.86	0.34	48,93,111,112	2
3	EOH	AAA	512	3/3	0.89	0.13	85,86,87,88	0
2	GOL	AAA	502	6/6	0.89	0.16	48,79,96,99	2

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 RF8 AAA 513:

$2mF_o$-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

There are no such residues in this entry.